MULTIDIMENSIONAL FREE ENERGY RELATIONSHIPS

IN ASYMMETRIC CATALYSIS

by

Kaid Clifton Harper

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

May 2013

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Kaid Clifton Harper**

has been approved by the following supervisory committee members:

Matthew S. Sigman	, Chair	**11-27-2012** Date Approved
Gary Keck	, Member	**11-27-2012** Date Approved
Ryan Looper	, Member	**11-27-2012** Date Approved
Joel Harris	, Member	**11-27-2012** Date Approved
Eric Schmidt	, Member	**11-27-2012** Date Approved

and by **Henry White**, Chair of the Department of **Chemistry**

and by Donna M. White, Interim Dean of The Graduate School.

ABSTRACT

Asymmetric catalysis is a powerful method for synthesizing enantiomerically enriched chiral building blocks. Detailed understanding of how catalysts impart facial bias on prochiral substrates has the potential to enable improved catalyst design and increase catalyst applicability. To this end, linear free energy relationships have been used to relate catalyst properties to enantioselectivity, enabling greater understanding of key catalyst-substrate interactions. Linear free energy relationships also can allow prediction of catalyst performance prior to their preparation. In this dissertation, several linear free energy relationships are described with a focus on developing predictive power and understanding the mechanism of asymmetric induction.

In asymmetric catalysis, steric effects are often implicated as key components in imparting enantioselectivity; however, they are typically treated empirically. In Chapter 2, steric parameters, particularly Charton parameters, are used to quantify ligand steric effects in the Nozaki-Hiyama-Kishi allylation of aryl aldehydes and ketones. Multidimensional linear free energy relationships, which simultaneously quantified the steric effects at both positions, are determined and used to predict ligand performance.

The multivariate linear free energy relationships have guided the design of a new ligand scaffold capable of enantioselective propargylation of ketones, which is discussed in Chapter 3.

The multivariate relationships were expanded to include nonsteric terms, which enabled the development of an electronically and sterically optimized catalytic system for the enantioselective propargylation of ketones, yielding enantioenriched homopropargyl alcohols.

The multivariate approach to describing substituent effects in asymmetric catalysis led to the evaluation of Sterimol parameters. Chapter 4 gives five examples of data sets where Sterimol values led to better correlation and predictive power than the previously used Charton parameters. The computational basis of the Sterimol parameters allows for greater interpretation of the models in which they are utilized.

Quantifying the factors that lead to enantioselective outcomes is a key challenge in asymmetric catalysis. Combining steric parameters, multidimensional analysis, and the principles of experimental design can lead to increased predictive power in asymmetric catalysis.

For my wife and parents.

TABLE OF CONTENTS

ABSTRACT iii

LIST OF ABBREVIATIONS viii

ACKNOWLEDGMENTS xiv

Chapter

1. LINEAR FREE ENERGY RELATIONSHIPS IN ASYMMETRIC CATALYSIS 1

Introduction 1
The Curtin-Hammett Principle 10
Ligand Electronic Effects in the Mn-catalyzed Epoxidation of *cis*-Alkenes 13
Catalyst Acidity in the Organocatalytic Hetero Diels-Alder Reaction 20
Polarizability in Thiourea Catalyzed Polyene Cyclization 25
Computed H-bond Length in the Asymmetric Strecker Reaction 29
Charton Steric Parameters in Asymmetric Nozaki-Hiyama-Kishi Allylation of Carbonyls 35
Conclusion 39
References 41

2. DEVELOPMENT OF STERIC-BASED THREE DIMENSIONAL LINEAR FREE ENERGY IN NOZAKI-HIYAMA-KISHI ALLYLATION REACTIONS 46

Introduction 46
Library Design and Synthesis 47
Developing a Model 53
The Principles of Experimental Design Applied to Asymmetric Catalysis 70
Reevaluation of the Data 75
Conclusion 81
Experimental 83
References 111

3. 3D FREE ENERGY RELATIONSHIPS AND THE PROPARGYLATION OF KETONES 114

Introduction 114
Library Evaluation of the NHK Propargylation of Ketones 117
Ligand Redesign 119
Synthesis of the Quinoline-Proline Ligand Library 123

Model Determination for the Quinoline-Proline Library 129
Conclusion 137
Experimental 138
References 165

4. MULTIDIMENSIONAL STERIC PARAMETERS IN THE ANALYSIS OF ASYMMETRIC CATALYTIC REACTIONS 170

Introduction 170
Comparison of Steric Parameters 171
Sterimol Parameters 175
Application of the Charton Parameter to Asymmetric Catalysis 177
Analysis of the NHK Allylation Reactions Using Sterimol Parameters 183
Evaluation of Substrate Steric Effects for the Desymmetrization of Bisphenols 187
Sterimol Analysis of the NHK Propargylation Reaction 194
Conclusion 199
Experimental 200
References 220

LIST OF ABBREVIATIONS

1-Ad	1-adamantyl
3D	three dimensional
Ac	acetyl
AcCl	acetyl chloride
Ac_2O	acetic anhydride
AcOH	acetic acid
AIBN	azobisisobutyronitrile
aq.	aqueous
Ar	aryl
atm	atmosphere
BINAP	2,2'-bis(diphenylphosphino)-1,1'-binaphthyl
BINOL	1,1'-Bi-2-naphthol
Boc	*tert*-butoxycarbonyl
Bn	benzyl
bs	broad singlet
BTM	benzotetramisole
Bu	butyl
*i*Bu	*iso*-butyl
IBCF	*iso*-butyl chloroformate
*t*Bu	*tert*-butyl

C	coefficient value matrix
°C	degrees Celsius
calcd	calculated
Cbz	carbobenzyloxy
$CHCl_3$	chloroform
cm	centimeter
CoMFA	comparative molecular filed analysis
Cr	chromium
Cy	cyclohexyl
d	doublet
Δ	heat
DAST	diethylaminosulfurtrifluoride
DCC	*N,N'*-dicyclohexylcarbodiimide
DCE	1,2-dichloroethane
DCM	dichloromethane
dd	doublet of doublets
ddd	doublet of doublet of doublets
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
dr	diastereomeric ratio
EDCI	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
ee	enantiomeric excess
equiv.	equivalents

er	enantiomeric ratio
ESI	electrospray ionization
Et	ethyl
Et_2O	diethyl ether
EtOAc	ethyl acetate
EtOH	ethanol
FTIR	fourier transform infrared spectroscopy
g	gram
GC	gas chromatography
h	hour
H-bond	hydrogen bond
hν	ultraviolet light
HDA	hetero-Diels-Alder
Hep	4-heptyl
HPLC	high pressure liquid chromatography
HRMS	high resolution mass spectrometry
Hz	Hertz
IBCF	*iso*-butylchloroformate
IR	infrared spectroscopy
K	equilibrium constant
k	rate constant
L	liter
LAH	lithium aluminum hydride
LFER	linear free energy relationship

LOO	leave-one-out
ln	natural logarithm
m	multiplet
M	molar
m	*meta*
Me	methyl
MeCN	acetonitrile
MEK	methyl ethyl ketone
MeOH	methanol
mg	milligram
MHz	megaHertz
min	minute
mL	milliliter
μL	microliter
mmol	millimole
μmol	micromole
mol	mole
MM	molecular mechanics
Mn	manganese
MP	melting point
MS	mass spectrometry
NBS	N-bromosuccinimide
NHK	Nozaki-Hiyama-Kishi
NMM	*N*-methylmorpholine

NMR	nuclear magnetic resonance
Nu	nucleophile
o	*ortho*
obsvd.	observed
OTf	trifluoromethylsulfonate
p	*para*
Pd/C	palladium on carbon
Ph	phenyl
PG	undefined protecting group
PMA	phosphomolybdic acid stain
ppm	parts per million
*i*Pr	*iso*-propyl
*i*PrOH	*iso*-propyl alcohol
q	quartet
QM	quantum mechanical
quinox	quinoline oxazoline
QuinPro	quinoline-proline
QSAR	quantitative structure-activity relationship
R	universal gas constant
RDS	rate determining step
R_f	retention factor
RT	room temperature
s	singlet or second
SFC	supercritical fluid chromatography

STD DEV	standard deviation
sub	substrate
T	temperature
t	triplet
TBAF	tetrabutylammonium fluoride
TBS	*tert*-butyldimethylsilyl
TEA	triethylamine
THF	tetrahydrofuran
TLC	thin layer chromatography
TMSCl	trimethylsilylchloride
tol	toluene
TsCl	*para*-toluenesulfonyl chloride
UV	ultraviolet
V	variance-covariance matrix
vs.	versus
X	design matrix
Y	response matrix

ACKNOWLEDGMENTS

Writing this dissertation has made me reflect on my entire education; in doing so, I have realized that all my achievements have been made possible by the help and sacrifice of many. First, I acknowledge and thank my advisor Matt Sigman for his mentorship and friendship. Over the past four years, he has taught me a great deal about many subjects ranging from organic chemistry to mentoring, and I appreciate his patience with me as I developed my abilities. I also appreciate the effort he puts forward in creating a group learning environment and high scientific standards.

I am thankful to past and present members of the Sigman group for their support and friendship. Specifically, I would like to thank Katrina Jensen for her kindness and also for her musings, which eventually turned into my initial project. I would like to thank Laura Steffens for her friendship from my first day of graduate school. I would most like to thank Ryan Deluca, Rachel Vaden and Benjamin Stokes, whose friendship over the years has made the lab one of the most enjoyable work experiences, even when facing difficulties. I am grateful to Elizabeth Bess for both her insight and collaboration on this project. I would also thank Katie Schafer, whose contributions as an undergraduate helped to define this work, and whose eventual friendship has eased the next transition in my life.

After 25 nearly continuous years of formal education, I have been the beneficiary of wonderful and dedicated teachers, and I thank all of them. From elementary to high school, I was continually challenged by these dedicated individuals who are too numerous to mention

specifically. From my undergraduate experience, I thank Professor Merritt Andrus and Mike Christiansen, who mentored me and helped develop my love of physical organic chemistry. From my graduate experience, I would also thank my committee members, each of which is a wonderful teacher. Specifically, I am grateful to Professor Joel Harris without whose help this project would not have taken shape.

Finally, I thank my family for their constant support and love. For my parents Kelly and Carilee, who are the greatest teachers I have ever known. All that I have that is good in me is from them. Most especially, I thank my wife, Jenny, whose constant love and support has allowed me to endure the difficult moments and rejoice in the happy moments. I am grateful for her sacrifices to allow me to pursue my education. I dedicate all of my work, past and future, to her.

CHAPTER 1

LINEAR FREE ENERGY RELATIONSHIPS IN ASYMMETRIC CATALYSIS

Introduction

Nobel Laureate William Knowles once wrote, "Since achieving 95% [enantiomeric excess] only involves energy differences of about 2 kcal [per mol], which is no more than the barrier encountered in a simple rotation of ethane, it is unlikely that before the fact one can predict what kind of ligand structures will be effective."[1] True to this statement, the field of asymmetric catalysis has relied on empiricism to develop catalysts. The challenge and the inherent interest in asymmetric catalysis is controlling the finite energies associated with asymmetric induction. This challenge is juxtaposed to the considerable potential that the enantioenriched products possess in synthetic chemistry. Even as the number of reported asymmetric catalytic methodologies expands at a rapid rate, the number of tools available to probe the origin of enantioselectivity in reactions remains relatively stagnant. However, if new reactions are to be applied broadly in organic synthesis, often additional studies must be performed beyond the initial report to ensure applicability to a desired substrate. Therefore, in general, synthetic chemists are not as willing to explore recently reported reactions where effort is required to generate the catalyst and the results are not certain or predictable. A modern goal in asymmetric catalysis is to provide a degree of predictability to these reaction outcomes and in so doing increase their implementation and acceptance in mainstream organic synthesis.

Prediction of enantioselective outcomes not only holds considerable potential in applying asymmetric reactions in synthetic chemistry but also presents a powerful tool in designing and optimizing catalysts within the field itself. Development of asymmetric catalysts is often a time consuming and difficult process. Two primary paths towards catalyst optimization have been reported in the literature.

The first is using a combinatorial chemistry approach by evaluating a large number of catalysts. This approach is often multigenerational with the highest performing catalysts becoming "parents" of further generations. A requirement in this approach is a modular ligand scaffold compatible with high throughput synthetic methods. The second path utilizes single catalyst probes designed to elucidate key elements responsible for enantioselectivity. This step-by-step approach is more hypothesis oriented but often requires a significant synthetic effort to create each meaningful catalyst probe. If the goal of developing new asymmetric catalytic reactions is simply generating high enantiomeric excess, at present there is no clear advantage in either path when more often than not the two paths converge.

In attempts to streamline and quantify catalyst performance, several techniques capable of predicting or correlating enantioselectivity have emerged. A purely computational approach has been applied to several systems, often with positive results.[2-9] A recent example from Houk and coworkers demonstrates how state-of-the-art computation techniques can be applied in organocatalysis.[10] They examined the benzotetramisole-catalyzed dynamic kinetic resolution of azlactones first reported by Birman and Li shown in Figure 1.1.[11,12] The resolution of azlactones is interesting synthetically because it provides direct access to enriched amino acid derivatives from racemic starting materials. Using a simplified system, B3LYP/6-31-G(d) calculations of the (*R*) and (*S*) transition states predicted only low enantioselectivity ($\Delta\Delta G^{\ddagger}$ = -0.2 kcal/mol) and opposite facial selectivity to that observed. To overcome the underestimated

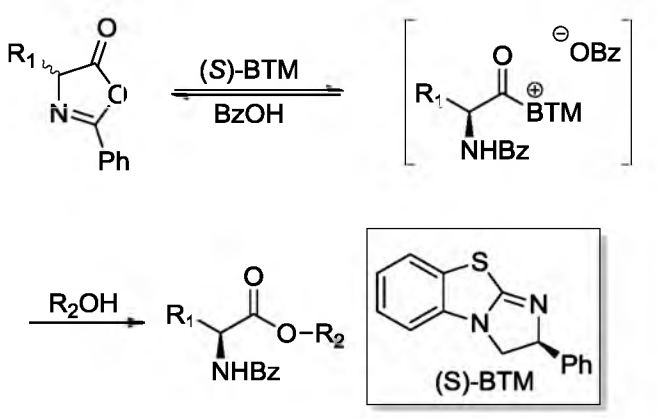


Figure 1.1. The benzotetramisole-catalyzed dynamic kinetic resolution of azlactones.

energy, Houk and coworkers turned to specialized functionals and hybrid basis sets shown in Table 1.1. The M06-2X/6-31G(d) geometry optimization gave the best results overall. M06-2X functional has been shown to model nonbonding interactions such as cation-π interactions, π-π interactions and hydrogen bonding with a higher degree of accuracy in several systems (*vida infra*).[13-15] With a reliable calculation of the transitions state structure, the key elements that imparted enantioselectivity became apparent. The cation-π interaction between catalyst and the alcohol nucleophile shown in Figure 1.2 rigidifies both diastereomeric transition states. The rate amplification responsible for resolution lies in the stabilization of the acyl anion by the benzoyl group in the transition state that leads to the major enantiomer (Figure 1.2). This type of carbonyl-carbonyl interaction is also found in crystal structures of proteins.[16-18] In the transition state leading to the minor enantiomer, the benzamide is orientated on the opposite side of the molecule and cannot stabilize the approach of the anion-nucleophile adduct.

The above example demonstrates that the B3LYP functional has limitations in estimating some types of key interactions present in catalytically relevant transition states.[2,3] In transition metal-mediated asymmetric catalysis, a greater challenge exists beyond selecting the correct functional. Molecular mechanics (MM) methods are often used to determine transition state structure where ground-state intermediates on either side of the transition state structure are well-defined energetically. MM methods require specific parameters, which are well-established for organic frameworks but do not exist for many transition metals. Thus, transition states involving metals are calculated using slower quantum mechanical (QM) methods. It is impractical to apply QM methods generally because they are time-consuming even for simple systems. Norrby and coworkers successfully hybridized the two types of calculation into a computational technique, which utilizes the speed of MM and the accuracy of QM in transition metal-mediated catalysis.[4,19,20] Essentially, they employ QM calculations of a simplified system

Table 1.1. Methods of calculation used to predict the enantioselectivity ($\Delta\Delta G^{\ddagger}$) in the kinetic resolution of azlactones. [Data from 10]

Method	$\Delta\Delta G^{\ddagger}$ (kcal/mol)
Experimental Data	1.2
B3LYP/6-31G(d)	-0.2
M06-2X/6-311+G (d,p)// B3LYP/6-31G(d)	1.9
B3LYP-D3/6-31G (d)// B3LYP/6-31G(d)	2.6
MP2/6-311+G (d)// B3LYP/6-31G(d)	2.1
M06-2X/6-31G(d)	1.8
M06-2X/6-311+G(d,p)// M06-2X/6-31G(d)	1.2
MP2/6-31g(d)// M06-2X/6-31G(d)	1.3

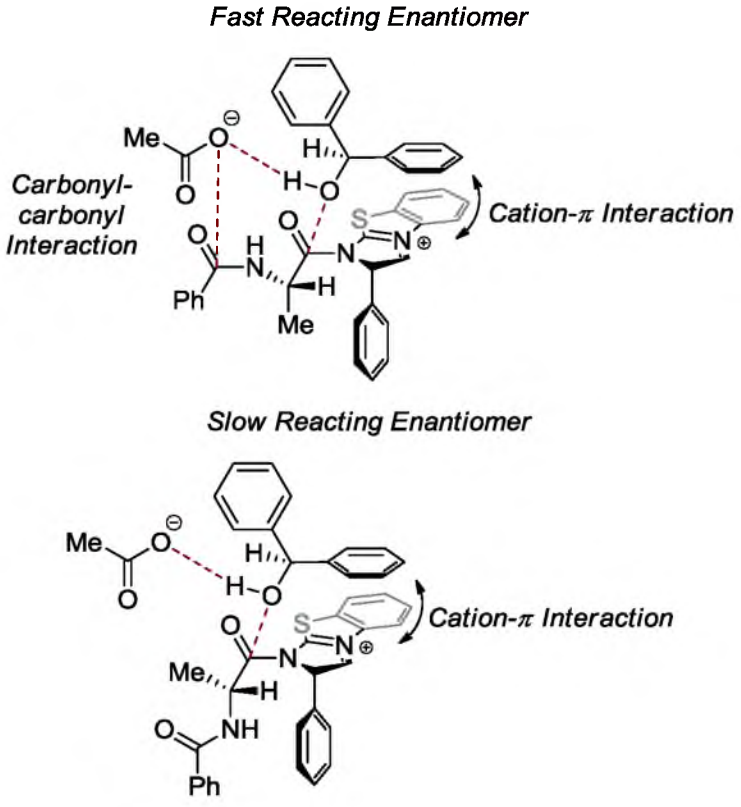


Figure 1.2. Calculated transitions states in kinetic resolution of azlactones showing the presence of a proposed stabilizing carbonyl-carbonyl interaction leading to high enantiomeric excess.

to develop a set of reaction specific parameters employed in a transition state force field. This force field can then be manipulated using faster MM methods and applied, evaluated, and optimized for enantioselectivity. Initially, Norrby and coworkers developed this model with the Os-catalyzed dihydroxylation reaction.[19,21-25] After benchmarking their techniques against this reaction, they have since applied it to rationalize the diastereoselectivity of dialkylzinc additions to aldehydes, Ag-catalyzed hydroamination of alkenes and the enantioselectivity of Rh-catalyzed hydrogenations successfully.[5,26,27] In the report of asymmetric hydrogenation, they were able to predict the enantioselectivities for a variety of ligand scaffolds and ligand/substrate pairings shown in Figure 1.3. This example highlights the advantages and potentially the future of in silico design of asymmetric catalysts.

These examples reveal that the viability of a completely in silico approach to asymmetric catalyst design depends on the ability to correctly determine the structure of the transition states. The key to accurately predicting enantioselectivity lies in the ability to precisely model the key interactions in not one but at least two diastereomeric transitions states. As basis sets advance their ability to accurately model complex interactions, the prediction of enantioselectivity completely in silico will become more of a reality. At the present time, a degree of uncertainty remains in transition state calculation and these calculations are usually verified by one of the few physical organic techniques available to probe transition state structure.

Efforts to boost the viability of computational techniques have coupled physical data with computation.[28-36] An interesting technique was reported by Denmark and coworkers where they applied the principles of quantitative structure activity relationships (QSAR), a widely applied method in drug design.[35] They investigated an asymmetric alkylation reaction using chiral ammonium ions developed in their lab and shown in Figure 1.4. Examining the

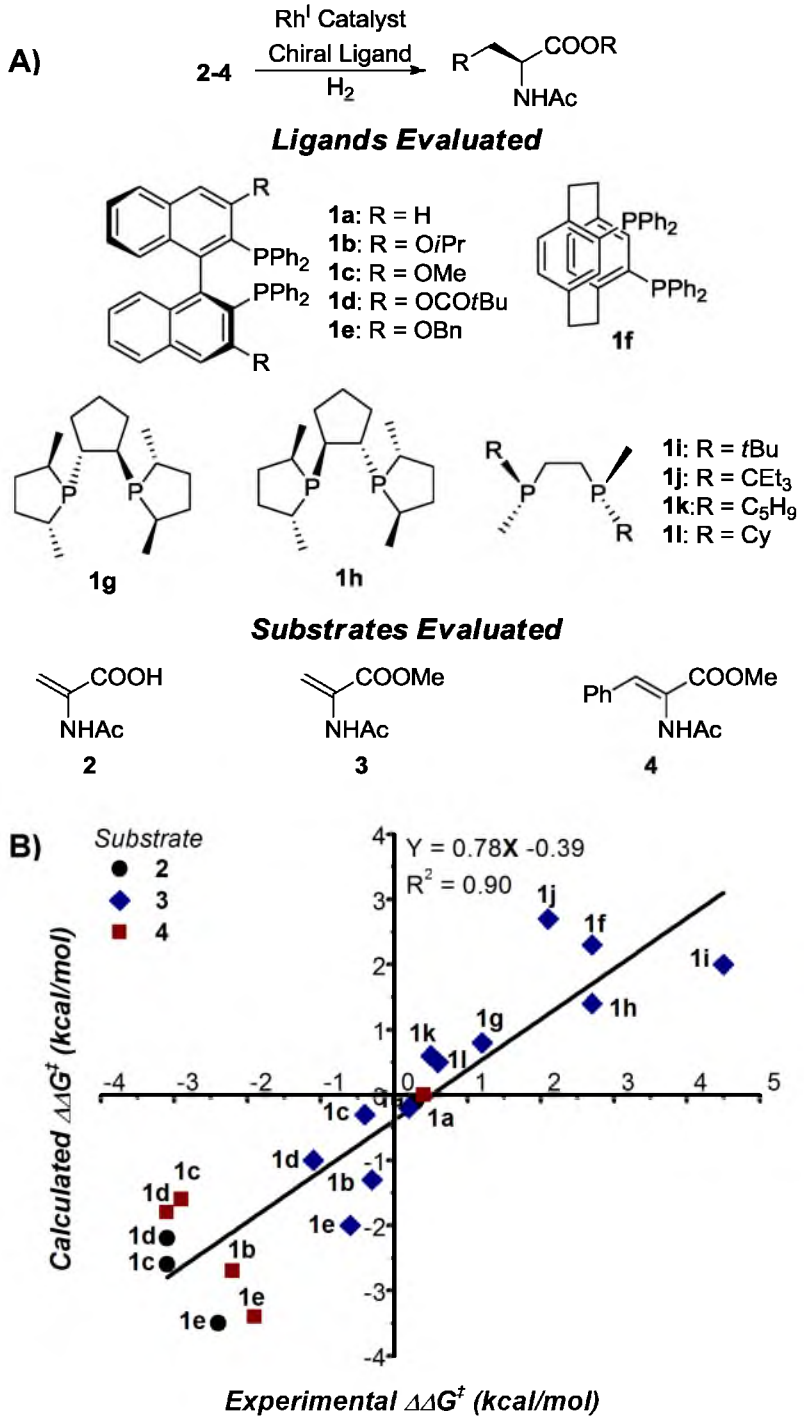


Figure 1.3. Predication and evaluation of several ligand and substrate pairings. A) The ligands and substrates explored experimentally and through calculation in the Rh-catalyzed asymmetric hydrogenation reaction. B) Plot of calculated enantioselectivity and experimentally observed enantioselectivity for the ligand substrate pairings shown in A. Deviation of the slope from unity represents calculated error. [Data from 27]

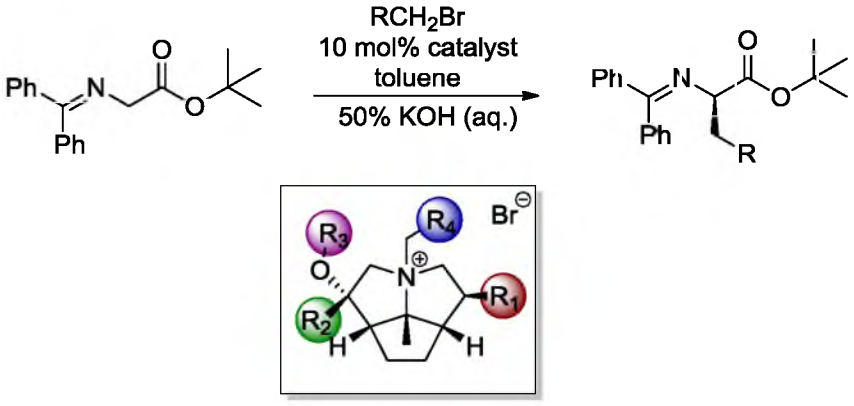


Figure 1.4. The asymmetric alkylation of glycine derivatives under phase-transfer conditions using a modular tetraalkylammonium salt.

enantioselectivity of the reaction, they employed comparative molecular field analysis (CoMFA).[37] CoMFA is considered a molecular interaction field technique where a molecule is described in steric and electronic elements. This is accomplished by encasing the molecule in a fine grid and interrogating it at each gridpoint with a point charge and an atom. The point charge interrogation provides information about the electronic nature of the molecule while the atomic interrogation provides steric information. CoMFA specifically refers to this gridpoint analysis when performed on a series of molecules that are related through a common factor. Prior to analysis, each catalyst's low energy conformation is determined through MM calculation and then aligned to a common element or framework. Denmark and coworkers used partial least squares regression to correlate the steric and electronic components in CoMFA with enantioselectivity, and identified regions where catalyst variation leads to greater enantioselectivity. Superimposing a proposed catalyst-substrate binding motif onto a visualization of these effects led them to formulate hypotheses about the source of enantioselectivity, which included several steric interactions as well as a proposed π-π stacking electronic interaction.

This technique, as well as others related to it, can provide powerful insight into the source of enantioselectivity and lead to improved design and understanding. The drawback of such an approach is that the synthetic effort required to generate libraries capable of interrogating simple systems is significant. Also, the active catalyst structure is assumed to be related to the ground state structure, which may not be a good assumption in asymmetric catalysis. Denmark and coworkers' system benefitted from a high degree of structural rigidity, which limited the number of potential conformers; nevertheless, it was challenging to accurately assess conformational effects on the model. In systems with a larger number of potential conformers, this problem could be compounded.

The computational methods outlined above represent a few select examples of the state-of-the-art in predicting and understanding the origins of asymmetric induction in catalytic reactions. The Sigman lab has pursued an alternative approach to attempt to predict and understand catalytic enantioselective reactions. Our and others' approaches have utilized linear free energy relationships (LFERs) to correlate substituent effects of the catalyst and substrate to enantioselectivity. Our interest in developing LFERs in enantioselective catalytic systems is based on the fundamental curiosity of how these reactions operate. Application of LFERs in asymmetric catalysis has closely paralleled quantitative structure activity relationship (QSAR) methodology, which has revolutionized medicinal chemistry.[38] In the process of our studies, we have become keenly aware of the techniques employed in the QSAR field and have endeavored to utilize this technology for developing our own studies. Whether or not LFERs possess the same potential to elucidate and manipulate key features in asymmetric catalysis remains to be seen and is the subject of this thesis. This chapter will examine several different reported LFERs, which correlate various catalyst properties to enantioselectivity. These studies have been instrumental in evolving the current design-based approach used for developing asymmetric catalysts in the Sigman laboratory.

The Curtin-Hammett Principle

LFERs have been used in physical organic chemistry for many years to probe reaction mechanism. The basis for all LFERs is found in the relation of the relative rate constant (k_{rel}) to the difference in the free energy of the transition state shown in Equation 1.1.[39]

$$\Delta G^{\ddagger} = -RT ln(k_{rel}) \quad (1.1)$$

Traditional LFER analysis used in physical organic chemistry has focused on defining and examining the rate-determining step of a reaction. Identifying and understanding the transition

state structure in the rate-determining step has led to countless catalyst improvements. In order to apply Equation 1.1 to asymmetric catalysis or any stereoselective reaction, the Curtin-Hammett principle must be applied.

The simplest interpretation of the Curtin-Hammett principle dictates that in reactions where there are multiple interconverting reaction isomers, conformers, or intermediates leading to a distribution of products, the distribution of products is principally determined by the largest energy barrier along the reaction coordinate, not necessarily the energy of interconversion between isomers, conformers or intermediates (Figure 1.5).[40,41] This principle applied to asymmetric catalysis relates to catalyst-substrate interactions through the enantiodetermining step and their effects on enantioselectivity. A simple example is a substrate binding to an asymmetric catalyst. The substrate can bind to the catalyst through a considerable number of conformers. Assuming substrate-catalyst binding is reversible and the barrier to intercoversion is low with regards to the enantiodetermining step, the product ratio or enantioselectivity observed is attributed solely to the difference in energy of the diastereomeric transition states and not the populated conformational states (Figure 1.5). Halpern and coworkers pioneered the application of the Curtin-Hammett principle in asymmetric catalysis, and they observed that the nature of catalyst-substrate interactions is normally assumed to be much weaker than the bond formation or cleavage occurring in reactions, indicating that the Curtin-Hammett principle should be applied with caution.[42] The application of the Curtin-Hammett principle allows the enantiomeric ratio (er), as determined typically by chiral separation, to be treated as a relative rate of formation of the two enantiomers. Using er as a relative rate also requires that the thermodynamic quantity be modified to reflect the relative difference in Gibb's free energy of the transition state or $\Delta\Delta G^{\ddagger}$ (Figure 1.5).

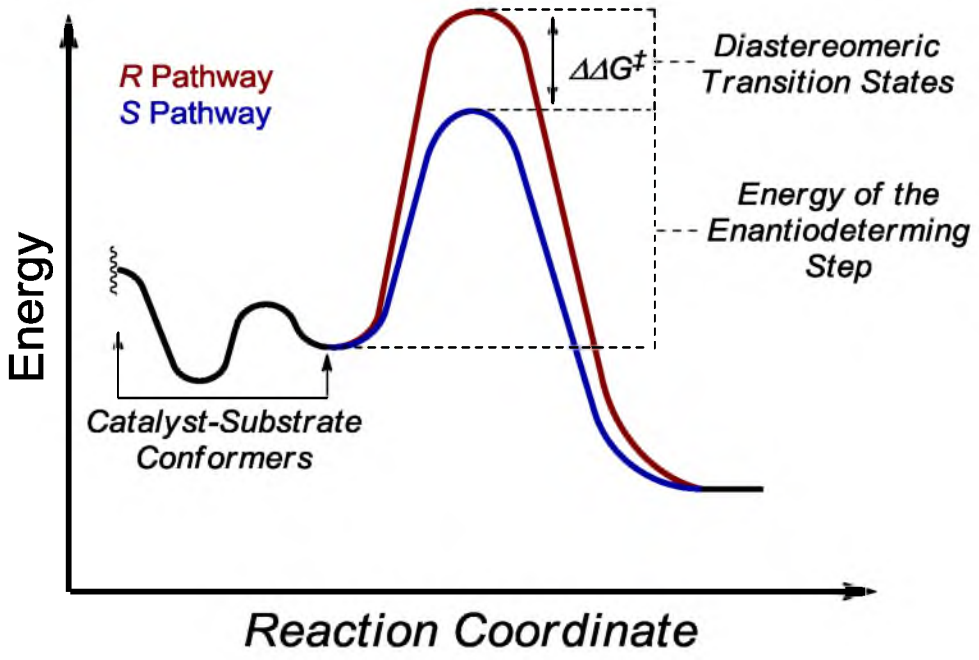


Figure 1.5. A catalytic asymmetric reaction coordinate that demonstrates the key features of the Curtin-Hammett principle.

$\Delta\Delta G^{\ddagger}$ is only a curiosity when considered as a single measured enantioselectivity, but examining $\Delta\Delta G^{\ddagger}$ as function of catalyst perturbation provides a glimpse into the key features of the transition state structure. The assumption in observing LFERs using a series of catalysts or substrates is that the mechanism of asymmetric induction is perturbed but not fundamentally changed. If the mechanism of asymmetric induction changes with variation of the catalyst, the comparability is lost. If a relationship can be drawn, it implies that the mechanism of asymmetric induction is robust to changes in the system.

The discussion so far has largely ignored one other key aspect of developing LFERs: identifying appropriate catalyst elements that can be systematically modified and quantified by parameterization. To accommodate systematic changes, a modular catalyst is ideal, as modularity increases the ability to explore a variety of potential effects. Furthermore, these potential effects must be accurately parameterized to encapsulate the properties of interest. The selected case studies in this chapter each represent novel application of different parameters in asymmetric catalysis and evaluate the information gained through correlative outcomes.

Ligand Electronic Effects in the Mn-catalyzed Epoxidation of *cis*-Alkenes

Although Halpern and coworkers first recognized the Curtin-Hammett principle in interpreting product enantiomeric ratio (er) in asymmetric catalysis, they did not report a LFER using this principle. The first and seminal report of a LFER in asymmetric catalysis was described by Jacobsen and coworkers in the context of the Mn-salen-catalyzed asymmetric epoxidation of *cis*-alkenes, a key method to access unfuctionalized chiral epoxides.[43] Mn-salen-catalyzed epoxidation is proposed to proceed via oxidation of Mn(III) to Mn(V), for which a number of suitable terminal oxidants have been reported (Figure 1.6). The resulting Mn(V)-oxo species

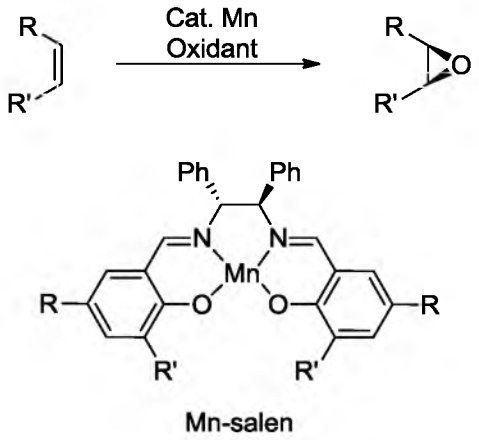


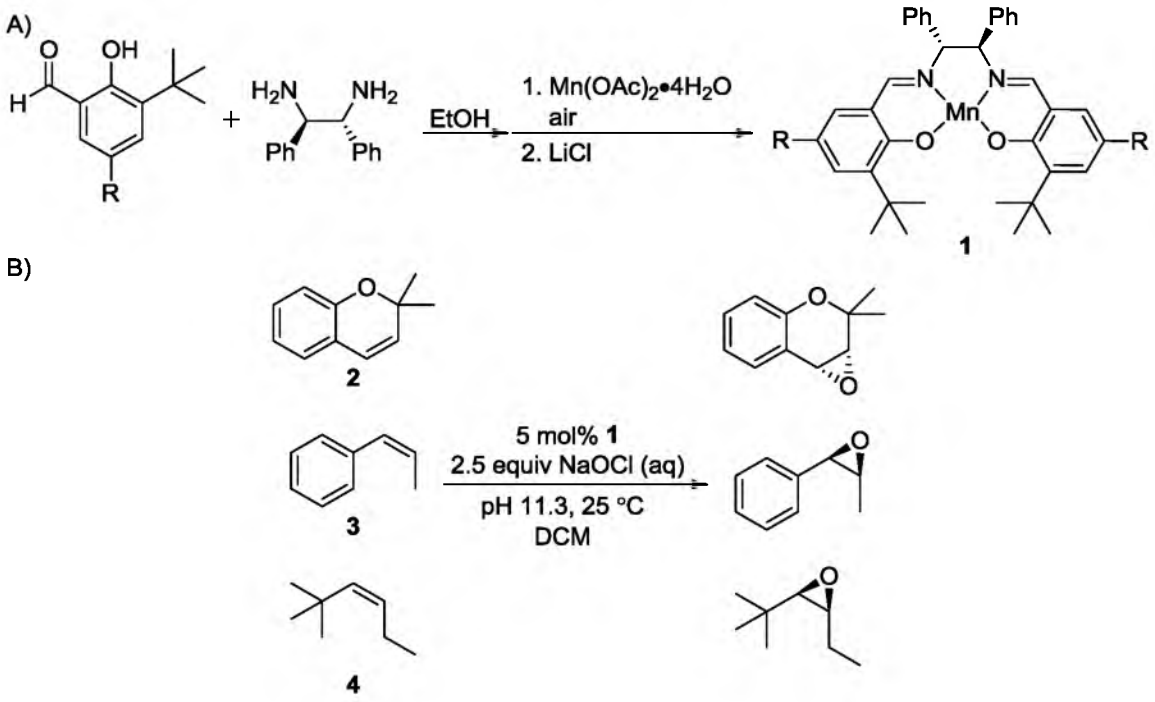
Figure 1.6. The manganese mediated epoxidation of alkenes.

readily reacts with various olefins, likely via a radical process, although debate still exists. For the reaction detailed below, aqueous bleach was used as the terminal oxidant, and Jacobsen and coworkers observed that the phase-transfer dynamics of Mn-oxidation were rate-limiting.[44]

The Jacobsen epoxidation was uniquely qualified for LFER anaylsis as the salen ligand template is modularly synthesized from readily available salicylaldehyde derivatives and a chiral diamine backbone (Figure 1.7A).[45] Catalyst assessment revealed a correlation between ligand electronic variation and enantioselectivity.[46] To quantify this electronic effect, σ-Hammett parameters, derived from the acidities of benzoic acid derivatives, were employed.[47]

Cis-2,2-dimethylchromene, *cis*-β-methylstyrene, and *cis*-2,2-dimethyl-3-hexene (Figure 1.7B-C) were all separately evaluated and each revealed a LFER with catalyst electronic nature as shown in Figure 1.8. The same general trend was observed for each substrate with electron-donating salens yielding the highest enantioselectivity. The sensitivities toward the electronic nature of the catalyst varied by substrate with *cis*-2,2-dimethylchromene displaying the greatest sensitivity and *cis*-2,2-dimethyl-3-hexene being the least sensitive.

To explain these observations, Jacobsen and coworkers invoked the Hammond postulate and hypothesized that the nature of the electronic effect was through bias for a more product-like transition state. The best evidence of this hypothesis is found in the slope of the Hammett plots. For each substrate, electron-donating ligands positively impact enantioselectivity, which is thought to originate by forming a more stabilized Mn(V)-oxo species, effectively making it a weaker oxidant and decreasing the rate of epoxidation. However, the weaker oxidant requires greater proximity of the alkene substrate in order for the epoxidation to occur. The increased proximity can in turn lead to greater substrate catalyst interaction, particularly, steric interactions between the chiral backbone of the salen ligand and the sterically dominate element of the alkene shown in Figure 1.9. These hypotheses were further



C)

Catalyst	*Substrate 2* er	*Substrate 3* er	*Substrate 4* er
1a: R = OMe	98:2	91.5:8.5	68.5:31.5
1b: R = Me	97:3	90.5:9.5	68.5:31.5
1c: R = H	94.5:5.5	89:11	67:33
1d: R = Cl	93:7	84:16	66:34
1e: R = NO_2	61:39	74.5:25.5	63:37

Figure 1.7. Electronic effects in the Jacobsen epoxidation. A) Synthesis of salen-type ligands. B) Substrates and conditions used to develop the ligand-based electronic LFER. C) Enantiomeric Ration (er) for multiple ligands and substrates. [45,46]

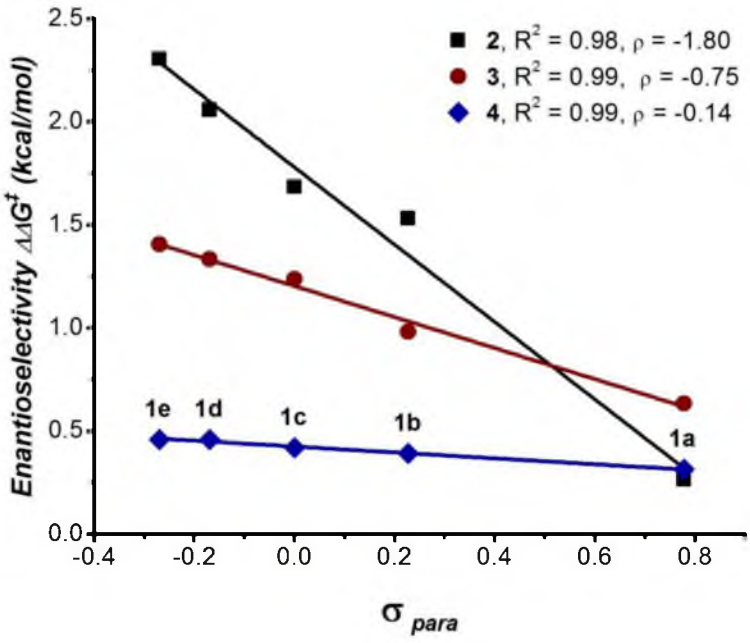


Figure 1.8. Plot of the enantioselectivities of substrates **2-4** with the Hammett σ values for catalysts **1a-1e**. [Data from 45,46]

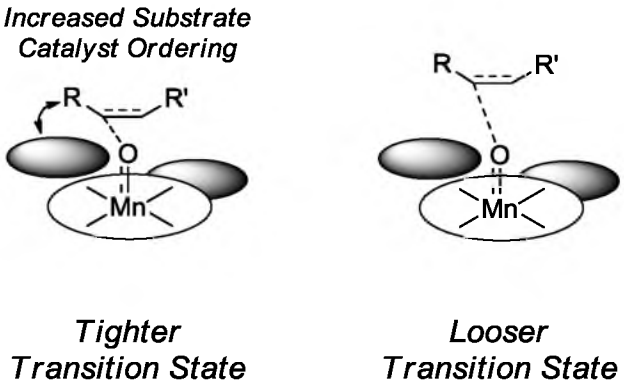


Figure 1.9. The hypothesized electronic effect on transition state structure results in varying degrees of tightness in the transition state and corresponding levels of enantioselectivity.

substantiated through kinetic isotope effects, Eyring analysis, and computational studies, all of which indicated that the electronic effect created a more product-like transition state.[48] Also supporting this hypothesis was a more recent study reported by Pericas and coworkers in which they examined the role of substrate electronics under modified epoxidation conditions using the commercial Jacobsen catalyst **5**.[49] They found a strong correlation between substrate electronics and enantioselectivity using trisubstituted olefins (Figure 1.10). They found the same trend originally observed by Jacobsen and coworkers for the electronic nature of the catalyst was mirrored in the substrate. Electron-rich alkenes, which are more reactive towards oxidation by the Mn(V)-oxo species, gave lower enantioselectivities, while electron-poor alkenes gave higher enantioselectivities. Comparison of the p values between the catalyst LFER and the substrate LFER reveals that the reaction is less sensitive to substrate electronics; however, this comparison may not be direct. The effect measured by Pericas and coworkers is for a highly conjugated system wherein the electronic perturbations would be delocalized over the alkene and the adjacent aryl rings, which would mitigate the electronic variation.

Jacobsen and coworkers' report introduced the field to the potential that LFERs have in asymmetric catalysis. The application of LFER analysis in this example elucidated a key nonintuitive interaction. Prior to this report, enantioselective outcomes had been primarily rationalized through steric effects. Jacobsen and coworkers' study and follow-up studies revealed electronic effects as important considerations in asymmetric catalysis. The information provided by the LFER about transition state structure led to a model for asymmetric induction, which has withstood multiple probes over a span of 20 years.

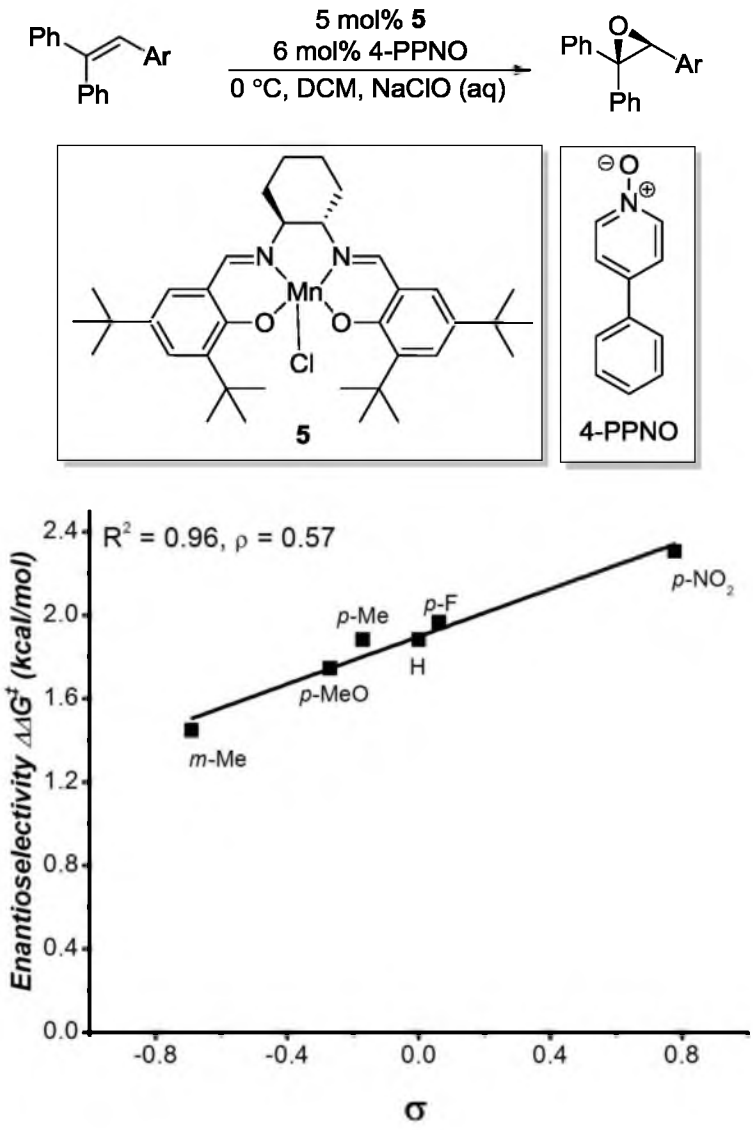


Figure 1.10. The experimental model reported by Pericas and coworkers evaluating the asymmetric epoxidation of trisubstituted alkenes with Jacobsen's catalyst. Plot of enantioselectivity as a function of substrate electronics in the same reaction. [Data from 49]

Catalyst Acidity in the Organocatalytic Hetero Diels-Alder Reaction

Jensen and Sigman employed the power of LFER analysis in asymmetric catalysis to a variant of the enantioselective hetero-Diels-Alder (HDA) reaction first reported by Rawal and coworkers.[50] After Rawal and coworkers initial report, they subsequently reported the rate enhancement and ultimately asymmetric catalysis by chiral α,α,α′,α′-tetra-2-napthyl-1,3-dioxolan-4,5-dimethanol.[51-54] The reaction was developed in a synthetic context, as the pyrone products generated through HDA are a common motif in natural products and active pharmaceuticals.

Sigman and Jensen became interested in the reaction to showcase a modular catalyst scaffold designed to be capable of H-bond catalysis.[55,56] The catalyst contains an oxazoline core to which is appended a serine derived tertiary alcohol and another amino acid derived amine (Figure 1.11). The modular nature of the catalyst allowed a wide variety of catalysts to be evaluated from readily available building blocks. Evaluation of a number of catalysts revealed the camphorsulfonamide derived catalyst **6** generated the highest enantioselectivity (Figure 1.12). Substitution of the camphorsulfonamide for a variety of amides revealed a surprising trend. Having initially assumed the high enantioselectivity exhibited by **6** was due to the bulky nature of the camphor appendage, the authors were surprised to observe a pronounced effect on enantioselectivity by simple amide groups.[57] Catalyst series **7a**-**e** revealed that more acidic catalysts yielded higher enantioselectivities for the HDA reaction (Figure 1.13). To develop the linear free energy relationship, the pK_a's of the corresponding acetic acids as measured in H_20 were employed (Figure 1.14). Brønsted acid-based LFERs have used acid pK_a's to correlate rate and acid catalysis in tradition physical organic chemistry for years.[39] The authors' use of pK_a's by analogy assumes that the substituent effects would scale similarly between acids and amides. The resulting LFER verifies this assumption, but because the inherent differences in H-bonding

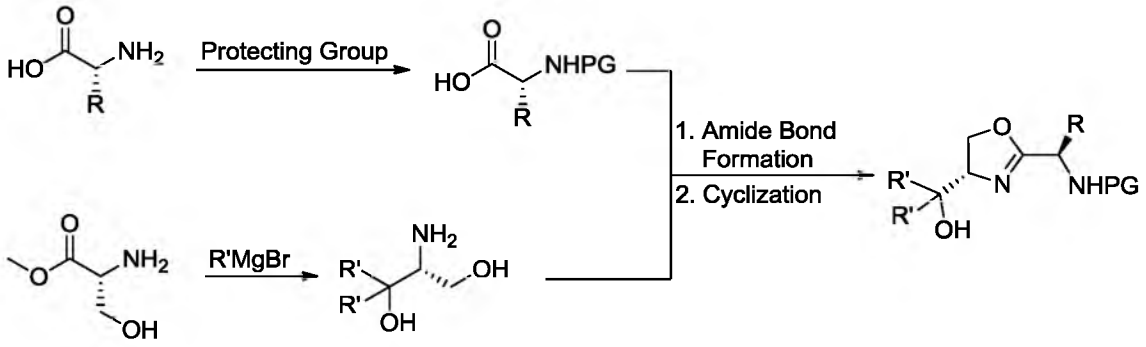


Figure 1.11. Modular synthesis of the H-bond catalyst framework.

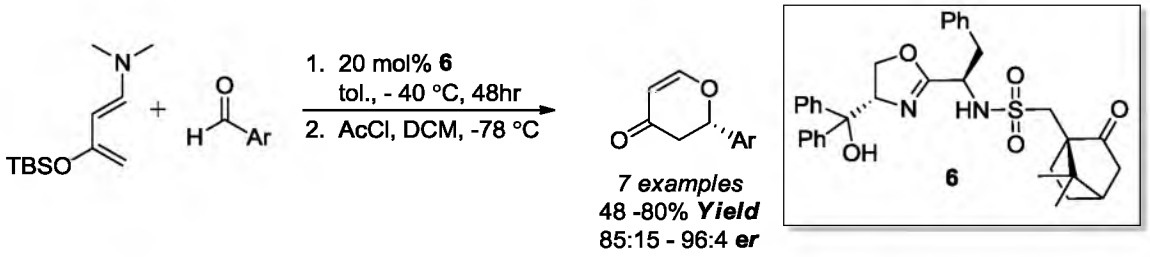


Figure 1.12. HDA reaction of Rawal's diene and aromatic aldehydes catalyzed by **6**. [57]

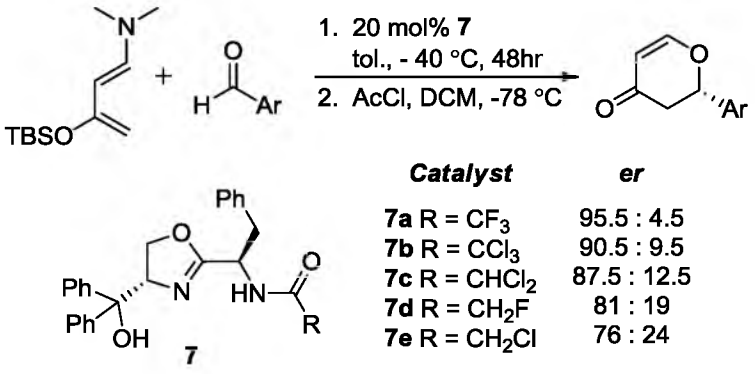


Figure 1.13. Evaluation of catalysts with different acidities in the asymmetric HDA reaction. [58]

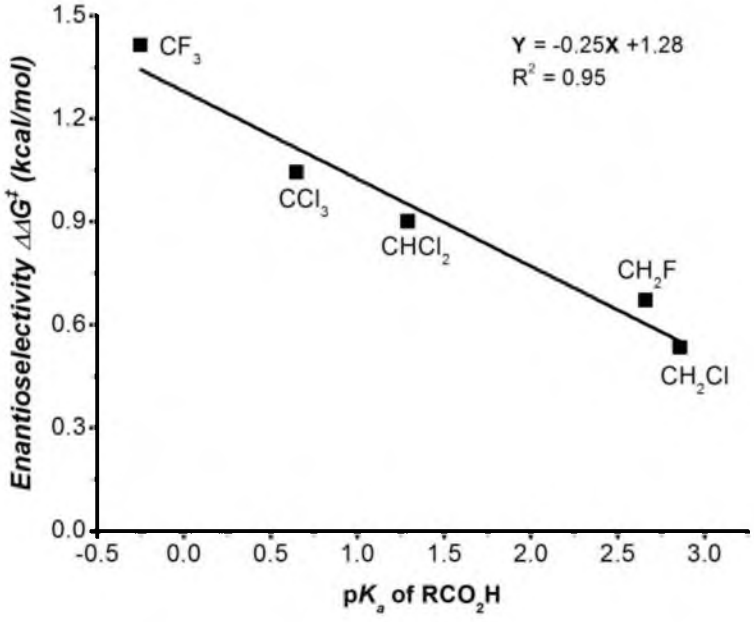


Figure 1.14. Plot of enantioselectivity as a function of catalyst acidity for the HDA reaction. [Data from 58]

and traditional acid catalysis, it remains to be seen if comparisons of the slope are relevant. This correlation implicates the strength of the H-bond formed between the substrate carbonyl, and the catalyst N-H bond is directly impacting enantioselectivity.

To fully explore the effect of amide N-H bond acidity on the system, a full kinetic study was undertaken.[58] The rate-determining step was shown to be the cycloaddition and not catalyst binding of the substrate. Kinetic data also suggested that the acidity of the catalyst affects the rate of substrate binding as well as the rate of reaction with diene. Similarly, the rate of formation of the major enantiomer was more sensitive to catalyst acidity.

To further examine the system, the authors exploited another modular aspect of the catalyst system, the substrate. Using a series of *para*-substituted benzaldehydes, a Hammett relationship was developed, which was predicted to mirror the effect of catalyst acidity (Figure 1.15). Evaluation of these substrates yielded no sensitivity between their electronic nature and the enantioselectivity of the reaction. However, a Hammett plot correlating substrate electronics and rate was observed at both low aldehyde and high aldehyde concentrations, which is consistent with the rate-determining step. At first glance, the strong correlation between catalyst acidity and enantioselectivity and the lack of correlation between substrate electronics and enantioselectivity is perplexing. If stronger H-bonding occurs as a result of effectively pairing pK_a's of the donor and acceptor, a relationship between substrate electronics and enantioselectivity would be expected.[59,60] Another hypothesis was formulated that explains the lack of substrate electronic effects via application of the Hammond postulate. Stronger catalyst acidity stabilizes the buildup of negative charge on the carbonyl oxygen creating a transition state where the substrate more closely resembles a product-like benzyl alcohol (Figure 1.16). The electronic substituent effects of benzyl alcohols have much less variation than the corresponding benzoic acids. The range of pK_a's of *para*-substituted benzyl alcohols is

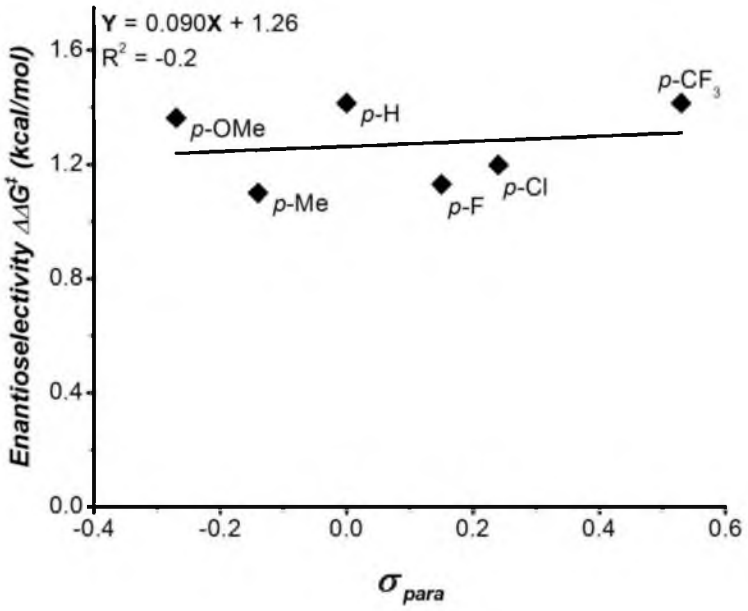


Figure 1.15. Plot of enantioselectivity as a function of substrate σ values for the HDA reaction. [Data from 58]

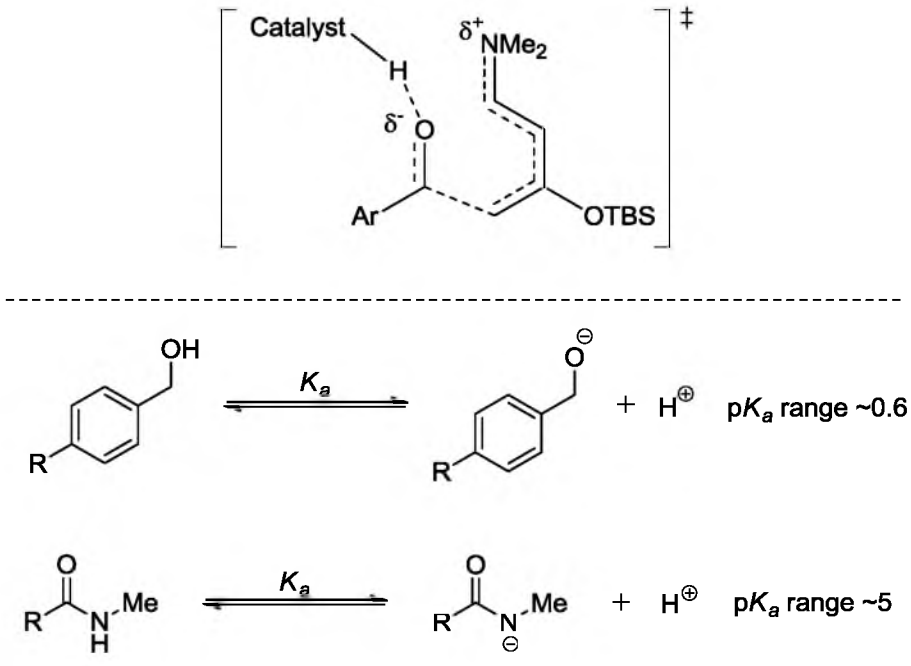


Figure 1.16. Proposed transition state for the HDA reaction which reflects more benzyl alcohol charater. The relative acidities of benzyl alcohol derivatives and amide derivatives.

roughly 0.6 pK_a units whereas the pK_a range of the substituted benzoic acids is 3.2; thus any lack of trend by the substrate could be easily attributed to experimental error.

In this case, the coupling of two LFER studies together with kinetic data provided the basis for a reasonable hypothesis of transition state interactions. The Brønsted-like correlation in an H-bond catalyzed system might find broader application as the number of enantioselective organocatalytic reactions that implicate H-bonding as a key element for enantioselection grows.

Polarizability in Thiourea Catalyzed Polyene Cyclization

Hydrogen bonding is a common motif for transition state stabilization in enzymes. Another common motif, which has come to light recently in understanding polyene cyclization, is cation-π interactions.[61,62] Cation-π interactions refer to the stabilization of cationic intermediates via electrostatic interaction with a π-system, typically an arene as demonstrated in Figure 1.17. This stabilization is facilitated by the polarizability of a molecule or its ability to disseminate charge.

Inspired by reports of these cation-π interactions in nature, Jacobsen and Knowles set out to design a catalyst capable of highly enantioselective polyene cyclizations.[63,64] The designed catalyst would combine the anion binding capabilities of thioureas as well as a moiety capable of stabilizing a cation via cation-π interaction (Figure 1.17). The result would be a catalyst capable of ionizing a substrate and providing a chiral environment for further reaction. The model reaction they studied was the bicyclization of hydroxyl lactams, which are known to ionize under acidic conditions (Figure 1.18). The N-acyliminum ion formed by ionization can be attacked by the nucleophilic alkene generating a carbocation, which can undergo another intramolecular addition by the arene.

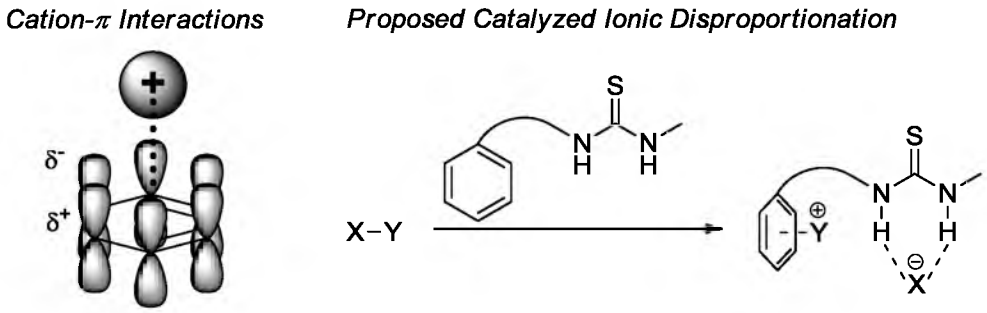
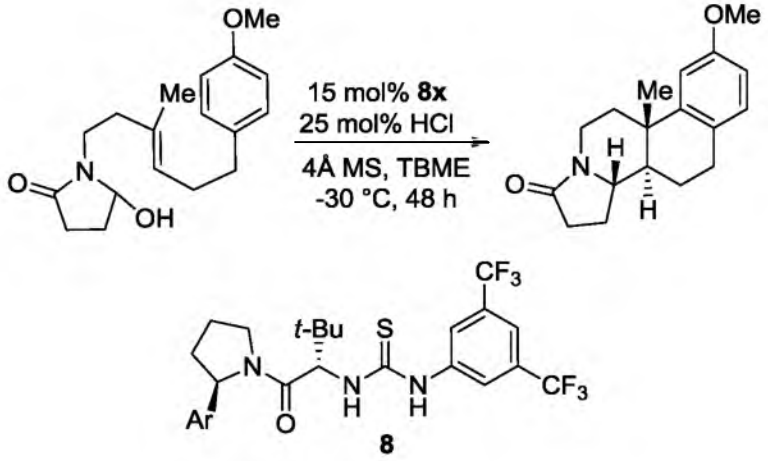


Figure 1.17. A cation-π interaction. Design elements of a catalyst capable of stabilizing ionic elements.



	Aryl =	er
8a	Phenyl	62.5 : 7.5
8b	2-Napthyl	80.5 : 19.5
8c	9-Phenanthryl	93.5 : 6.5
8d	4-Pyrenyl	97.5 : 2.5

Figure 1.18. The enantioselective cyclization of hydroxyl-lactams.

The catalyst designed made use of the well-characterized bistrifluoromethylphenyl thiourea employed by many in this field, connected by an amide linker to a chiral aryl pyrolidene. Proof of their design concept was exhibited by catalyst **8** in the model system, although in low yield and low enantioselectivity. Expanding the size of the aryl ring led to better yields and improved enantioselectivities. It should be noted that the reaction forms three new contiguous stereocenters through separate bond-forming events, and the reported enantioselectivities are for the single diastereomer formed in the reaction.

To determine the role of the arene, Jacobsen and Knowles correlated enantioselectivity with arene polarizability for catalysts **8a**-**8d** (Figure 1.19).[65] The measure of an arene's polarizability is its capability to delocalize charge through distortion. The correlation between polarizability and enantioselectivity implies that the catalysts were stabilizing the cationic intermediates by delocalizing positive charge.[66] The correlation indicates even larger aryl rings would generate higher selectivity; however, extrapolation of this LFER as a design principle was not explored. This might be due to the fact that polarizability values for larger substituents are not available, and the corresponding aryl bromides are not commercially available.

The LFER implicates the ability of the extended π-systems to stabilize cationic charge but did not rule out the argument that the aryl ring's effect is steric and not electronic in nature. To delineate the role of the arene, they evaluated the effect of temperature on enantioselectivity with each catalyst (Figure 1.20). The resultant Eyring analysis showed that varying the aryl ring had a primarily enthalpic effect. This is consistent with energetic stabilization of the cation intermediates, as such stabilization would be primarily enthalpic with a negligible entropic element.[67] Conversely, if the role of the aryl ring was primarily a steric effect, the Eyring analysis would have revealed an entropic effect relating to substrate ordering.

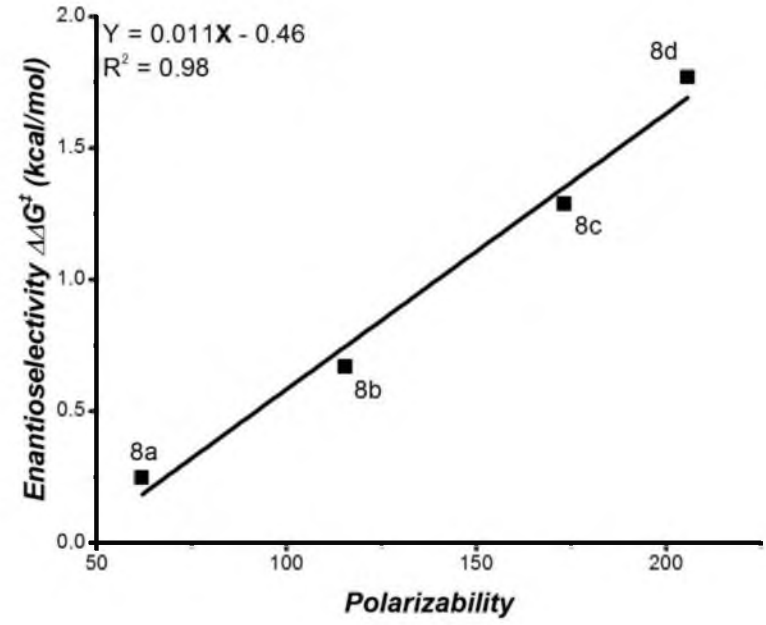


Figure 1.19. Plot of enantioselectivity as a function of arene polarizability in the asymmetric bicyclization of hydroxyl lactams. [Data from 63,64]

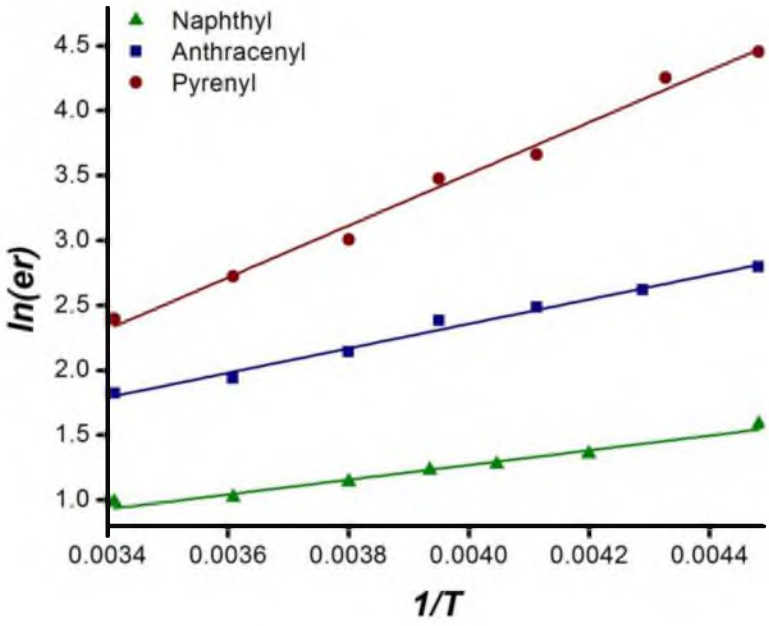


Figure 1.20. Eyring analysis comparing the roles of aryl substituents in the bicylcization of hydroxyl lactams. [Data from 63,64]

The mechanism of asymmetric induction in the polycyclization reaction is less clear. The reaction presumably proceeds through a closed six-membered transition state, which results in high diastereoselectivity. The catalyst might be capable of stabilizing each of three separate cations formed in the reaction pathway. The LFER indicates that the catalyst is interacting with the initial N-acylimminium ion to form the first chiral center, whether or not the catalyst remains in contact with the substrate after the initial enantioselective bond forming event is not clear. It seems reasonable that given the strong enthalpic contribution generated by the cation-π stabilization that the subsequent cations would remain in contact with the catalyst. However, the remaining bonds could be formed through favorable diastereoselective pathways.

This correlation between catalyst polarizability and enantioselectivity not only constitutes an important novel LFER with a noncovalent attractive interaction but quantifies an important design element in asymmetric catalysis. Although cation-π interactions might play significant roles in other reactions they had never been so directly implicated and quantified. This will be an important consideration in the future for reactions where cationic intermediates are accessible.

Computed H-bond Length in the Asymmetric Strecker Reaction

Among organocatalytic reactions, few have received as much attention as the enantioselective Strecker reaction.[68] In the Strecker reaction, nucleophilic cyanide is added to a an imine via 1,2-addition (Figure 1.21). The products of asymmetric Strecker reactions are synthetic precursors to many unnatural amino acids. Jacobsen and Sigman first reported an enantioselective variant of the Strecker reaction in the late 1990's.[69-71] The culmination of this work was reported in 2009 with a simplified highly enantioselective Strecker catalyst compatible with a wide range of substrates (Figure 1.21).[72] The reaction also used nonvolatile cyanide

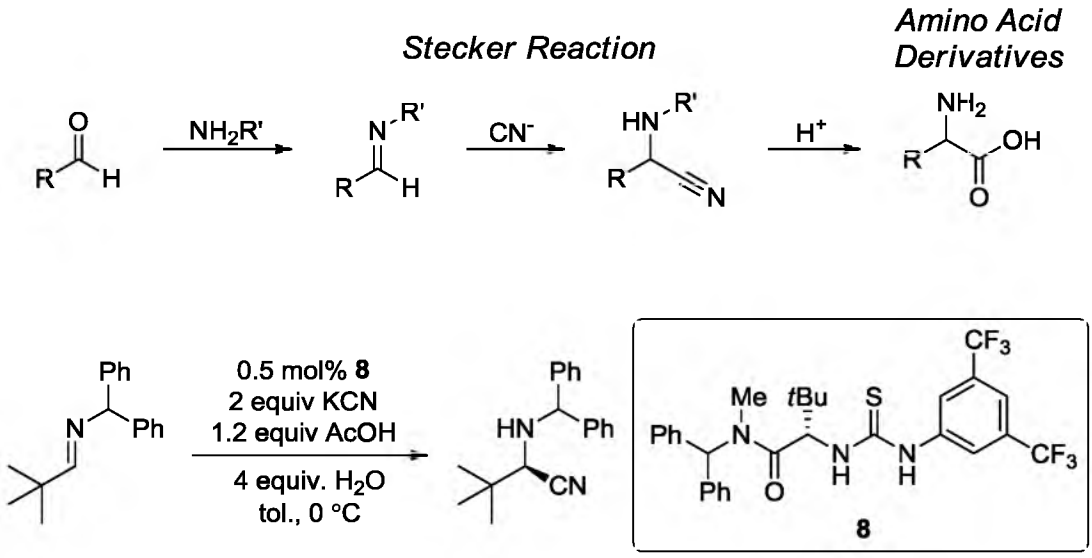


Figure 1.21. The Strecker reaction and hydrolysis to form amino acid derivatives. The asymmetric Strecker reaction for tertiary imines.

sources and moderate temperatures as compared to previous iterations. In an effort to understand the subtleties of this powerful reaction, Jacobsen and Zeund undertook a kinetic, physical organic, and computational study of their system.[14]

As a part of their kinetic and optimization studies, they generated a small library of thiourea based catalysts (Figure 1.22). Upon first inspection, these catalysts possess different properties and do not contain a complementary set of variations, as would be required to develop a traditional LFER. Using this set of catalysts, they computed the energy differences between the major and minor enantiomeric pathways at three different levels of theory, B3LYP/6-31G(d), M05-2X/6-31+G(d,p) and MP2/6-31G(d). In each case, correlation was found between the calculated $\Delta\Delta E^{\ddagger}$ and the observed $\Delta\Delta G^{\ddagger}$. Interestingly, B3LYP/6-31G(d) was shown to be the most accurate level of theory for the system, despite its propensity to underestimate the energies associated with noncovalent attractive interactions. Although the calculations consistently overestimate the $\Delta\Delta E^{\ddagger}$ values, the correlation to observed enantioselectivity suggests that the error is systematic. Also, the computation correctly predicted the growing energetic preference for the *R* enantiomer across the catalyst set. Considering the amount of variation within the catalyst library, the correlation verifies the viability of computation for examining the system.

Exploring the computed structure for each catalyst revealed no obvious steric interaction that could explain increased enantioselectivity. The spatial arrangement of atoms was either static through the series or deemed inconsequential to the enantioselective outcome. This observation raised the question of how the variation in enantioselectivity is achieved for the different catalysts. In fact, the calculations revealed no significant difference in the H-bond lengths between the (*R*) or (*S*) product forming pathways for highly or poorly enantioselective catalysts. However, their computational work had revealed that the computed

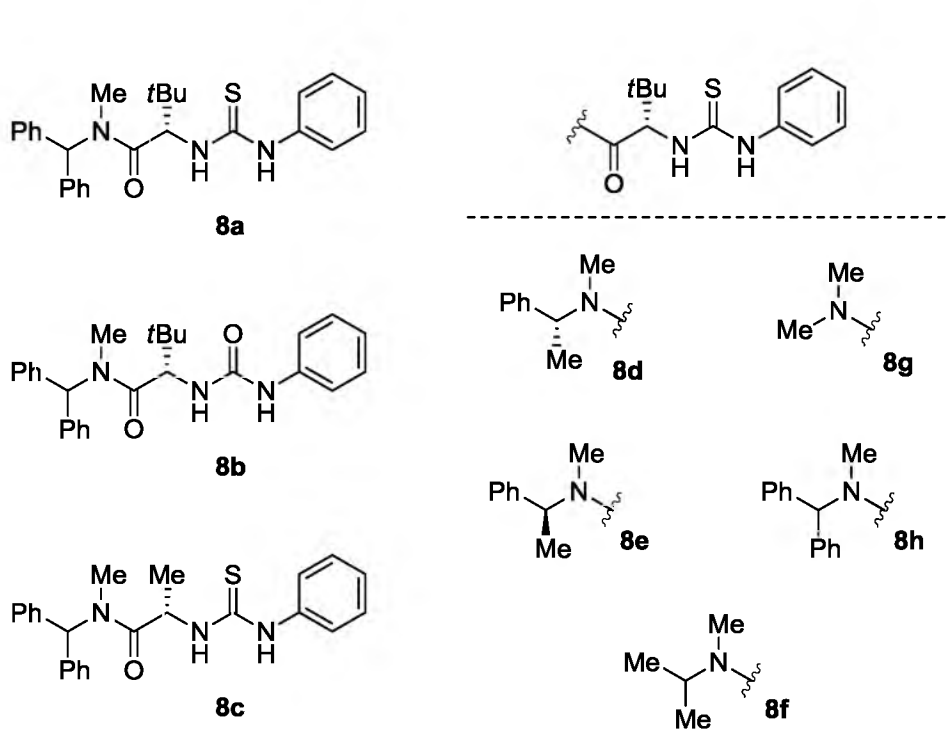


Figure 1.22. Library of catalysts used to evaluate the Strecker reaction computationally and experimentally.

rate-determining step of the reaction was rearrangement of the ion pair through a carbonyl stabilized H-bonding network (Figure 1.23A). They examined the role of this H-bond network through this step and identified no strong correlation between the cumulative H-bond distances in the *R* selective pathway. However, in the *S*-selective pathway they observed a LFER between cumulative H-bond distance and enantioselectivity (Figure 1.23B).

This LFER provides compelling evidence that the source of enantioselectivity is due to weaker stabilization of the imminium ion in the *S*-pathway. For the more enantioselective catalysts, the amide carbonyl becomes less accessible in the preferred transition state geometry inherent to the *S*-pathway which leads to its destabilization relative to the *R*-pathway. This highlights another feature of H-bonding not discussed previously: the directionality of the bond matters. In this system, there are no direct steric interactions that explain destabilization of a specific pathway. Instead, the steric effect arises from the catalyst itself, where its low energy conformation presumably leads to subtle differences in the amide carbonyl direction relative to the thiourea. This, in turn, leads to increased differences in the H-bonding network responsible for stabilization of the key intermediate.

This study presents another case where LFER analysis provided evidence for a nonintuitive catalyst-substrate interaction and implied transition state structure. Not only does it present LFER analysis in the development of an extremely powerful synthetic reaction, but it represents the melding of LFERs with computational chemistry. While evaluation and prediction of catalysts in silico has not yet fully arrived, this work presents an effective application of computational chemistry to develop a set of specialized parameters (cumulative H-bond length) and correlate them to experimentally derived results. Specifically, the use of such parameters in LFER analysis to evaluate or support specific mechanistic hypothesis has potential application in asymmetric catalysis.

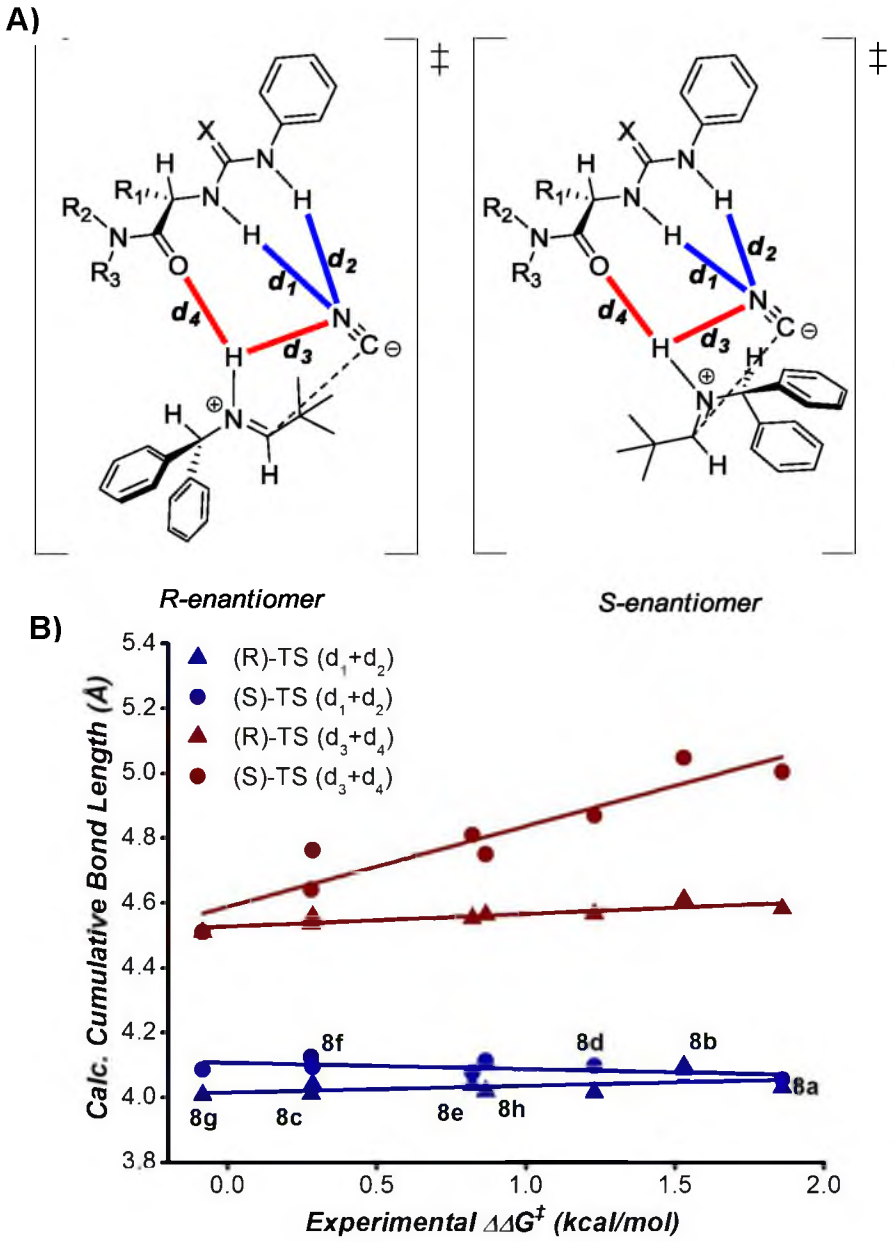


Figure 1.23. Prediction of enantioselectivities using H-bond length. A) Calculation of the proposed enantiodetermining transitions states leading to the *R* and *S* enantiomers (B3LYP-/6-31G(d)) The key bonding interactions are plotted and labeled. B) Plot of calculated bond lengths as function of experimentally observed enantioselectivity. [Data from 14]

Charton Steric Parameters in Asymmetric Nozaki-Hiyama-Kishi Allylation of Carbonyls

The previous example demonstrates how a subtle steric effect can have a profound influence on enantioselectivity. As previously stated, steric effects are widely implicated in asymmetric catalysis. Although several sets of experimentally based steric parameters have existed for years, no real effort to correlate steric effects to enantioselectivity existed until our group became interested in correlating a pronounced steric effect discovered in our investigations of the Cr-mediated Nozaki-Hiyama-Kishi (NHK) additions of allyl fragments to carbonyls.[73]

The NHK reaction mechanism is outlined in Figure 1.24 and involves addition of Cr(II) into an allylic bromide bond.[74] The subsequent Cr(III) allyl species is Lewis acidic and will activate a carbonyl to undergo nucleophilic addition from the pendant alkene. This nucleophilic addition occurs, presumably, through a closed, six-membered transition state. Initially reported as a reaction requiring super-stoichiometric amounts of Cr, it was rendered catalytic in Cr by addition of a terminal reductant for Cr(II), typically manganese, and a species capable of sequestering the Cr-alkoxide (TMSCl).[74] The reaction is just one of many 1,2-allylations of carbonyls, the synthetic utility of which is well-documented.

Sigman and Lee became interested in these reactions as a platform for the same amino acid-oxazoline ligand template as previously described.[75,76] Their initial entry was the report of ligand **9** imparting high degrees of enantioselectivity in the allylation of aryl aldehydes (Figure 1.25A). After catalyst modification, Sigman and Miller reported the expansion of the substrate scope to aryl ketones, for which no previous NHK asymmetric methodology had been reported (Figure 1.25B).[76] In many empirical iterations of the modular catalyst, a significant steric effect was observed through manipulation of the carbamoyl group. This observation, in combination

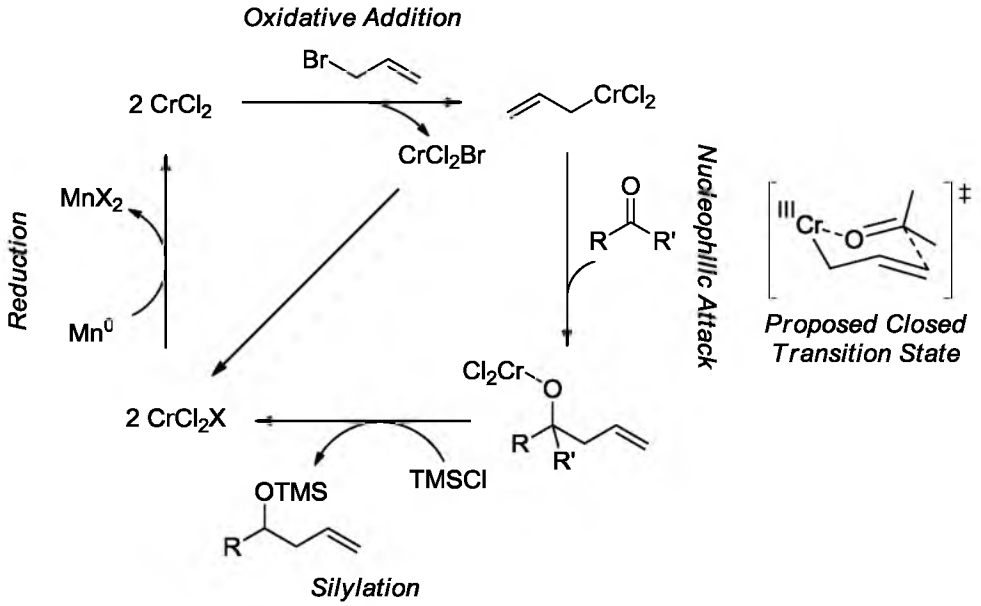


Figure 1.24. Proposed catalytic cycle for the Cr-mediated NHK allylation reaction with the proposed cyclic transition state shown.

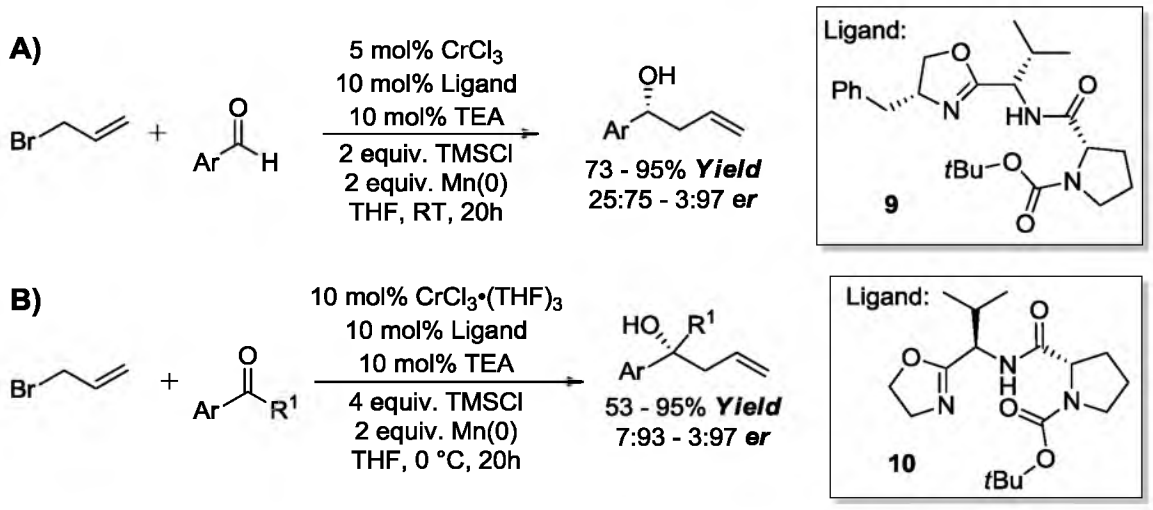


Figure 1.25. Previously reported reaction conditions for the NHK allylation of aldehydes and ketones. [72,73]

with the modular nature of the ligand template, provoked further investigation into the role of the carbamate.

In contrast to all of the previously discussed studies, this study of steric effects in the NHK reaction was not accompanied by a detailed examination of mechanism. Preliminary investigations into the reaction suggested that the rate-limiting step was silylation of the Cr-alkoxide and prior steps are relatively fast, a conclusion that is generally supported in the NHK literature. This suggests mechanistic and structural information cannot be obtained using kinetic analysis. In addition, the ground state catalyst has resisted any attempts to crystalize, providing the impetus to study the system using LFER analysis.

Regardless of this lack of general information, a model system was selected using catalyst framework **10** to examine two model reactions: the allylation of benzaldehyde and acetophenone (Figure 1.26).[73] Although ligand **10** was not an optimal ligand for ketone allylation, it had proven a competent ligand for the reaction. Variation of the carbamoyl group gave a series of ligands **10a-e**, which were evaluated in the model reactions. The results showed significant sensitivity to the presumed steric effects at this position.

The difficulty in quantifying and correlating this steric effect lies with the parameter choice. A significant portion of Chapter 4 will explore various steric parameters in depth and is beyond the scope of this introductory chapter. In short, Charton parameters were selected largely due to their larger number of reported values.[77-80] The nature of the Charton parameter is based on Taft's classical experiments to delineate steric effects from electronic effects (Figure 1.27).[81-83] Charton evaluated Taft's experimental data and found correlation between it and the calculated Van Der Waals radii of the substituent. The correlation allowed Charton to extrapolate Taft's data set, and generate parameters for a large number of substituents.

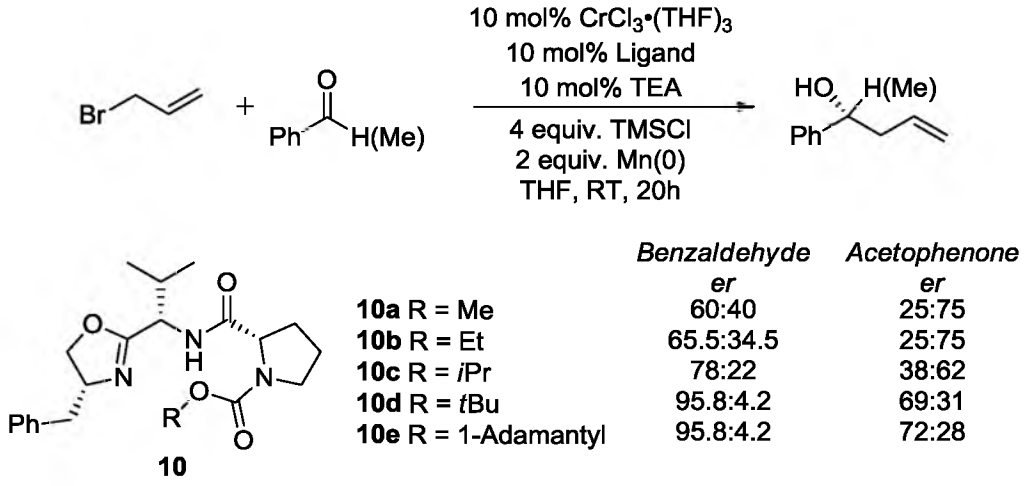


Figure 1.26. The exploration of steric effects in the NHK allylation. [Data from 73]

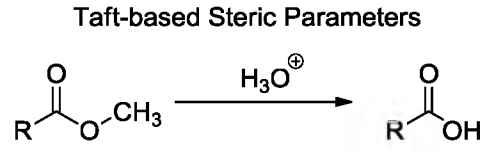


Figure 1.27. The experimental model used to derive Taft's steric parameters.

Applying Charton's parameters to the NHK allylation reaction led to a strong correlation between the substituents size and the enantioselectivity (Figure 1.28). The slopes also indicated that the reaction was very sensitive to steric bulk at the carbamate position. It was this sensitivity that was explored as a catalyst design element in extrapolation of the LFER. Using the available Charton parameters, three larger substituents were selected for incorporation into the ligand and subsequent evaluation in the NHK reaction. Using the LFER, the enantioselectivities for these substituents was predicted to be beyond the previously reported optimized system. However, evaluation of these catalysts manifested a break in the correlation and rendered the LFER ineffective as a predictive tool (Figure 1.29).[84] The resolution of this perceived break in linearity will be the focus of much of this dissertation.

The observation of this type of steric effect was not unique in asymmetric catalysis. Evaluation of published data revealed that Charton steric parameters could correlate steric effects in numerous systems.[84] In contrast to the above cases, the LFER was used as a design element more than a tool to derive the mechanism of asymmetric induction for the NHK allylation reactions. This study does represent the first successful attempt to correlate steric effects in asymmetric catalysis and reveals Charton parameters as potential tools to examine such effects. The application of steric parameters to asymmetric catalysis provided the impetus for what will be described in the remainder of this dissertation.

Conclusion

These examples showcase the power of LFERs in asymmetric catalysis. The use of LFERs can provide key insight into transition state structure and contribute to the analysis of the reaction. It is insight into these transition state structures that will prove vital to applying these reactions generally in organic synthesis. These examples are only a sampling of the various

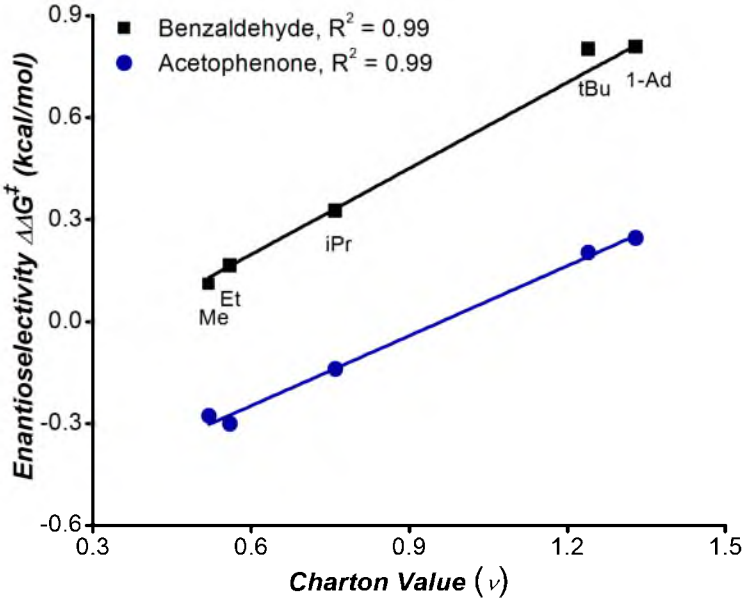


Figure 1.28. Plot of enantioselectivity of the NHK allylation reaction as a function of Charton's steric parameters. [Data from 73]

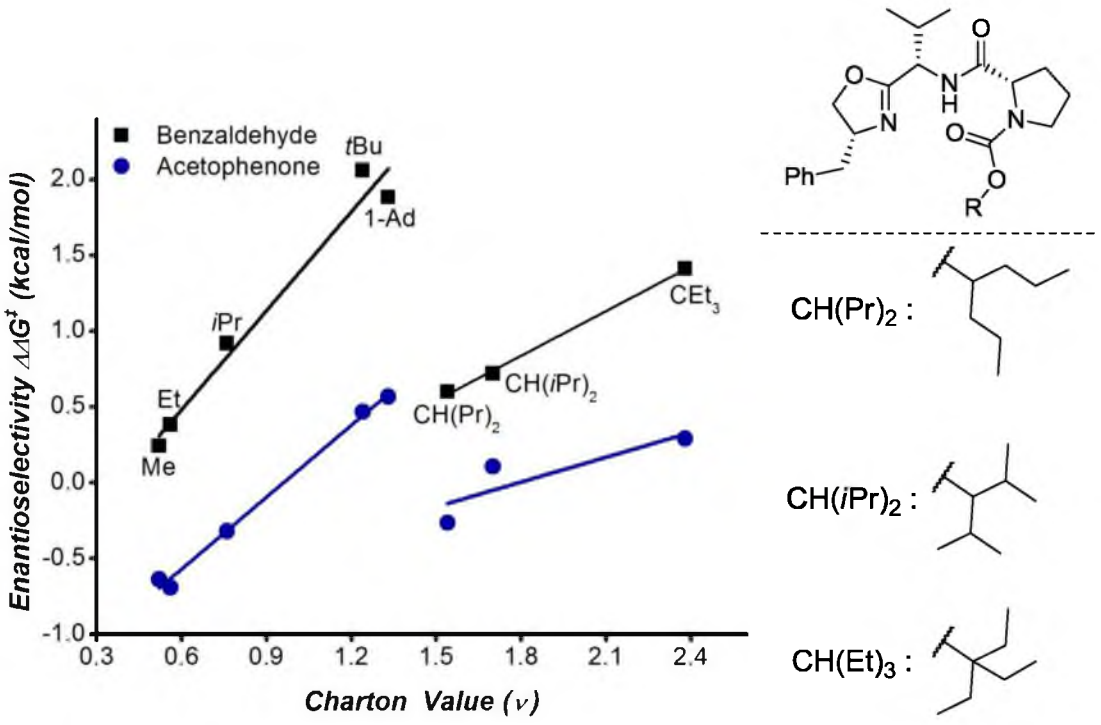


Figure 1.29. Plot of the enantioselectivity of the NHK allylation as a function of Charton steric parameters showing the nonlinear nature for larger substituents. [Data from 73]

substituent effects that can be explored through LFER analysis. The requirements of LFER analysis are that the system be robust to changes in the overall mechanism of asymmetric induction, the system must possess some degree of modularity, and parameters must exist that encapsulate the systematic changes and their effect. Development of LFERs requires synthetic effort to arrive at catalyst libraries. In this regard, LFERs are no different than many of the other tools available to probe asymmetric reactions. The element that sets LFERs apart is the wealth and character of information available should the analysis prove fruitful. LFER development can be complementary to computational-based designs, and the combination of the two is a powerful approach where computation can be used to arrive at unique parameters and correlated to enantioselectivity.

This chapter has focused on a handful of examples, in which a wide variety of parameters have been used to develop these LFERs. In order to maximize the information inferred from the experimentally based parameters, a larger number of well-understood reactions must be studied through LFER as well as analogous techniques to generate comparability. The variety of parameters discussed might imply that the field is rich with examples of LFERs but, regrettably, only a handful of other examples of LFERs in asymmetric catalysis have been reported.[85-87] Given the information attainable through such application, the hope is that this number will grow in the coming years.

References

(1) Knowles, W. S. *Acc. Chem. Res.* **1983**, *16*, 106.

(2) Cheong, P. H.-Y.; Legault, C. Y.; Um, J. M.; Çelebi-Ölçüm, N.; Houk, K. N. *Chem. Rev.* **2011**, *111*, 5042.

(3) Houk, K. N.; Cheong, P. H.-Y. *Nature* **2008**, *455*, 309.

(4) Donoghue, P. J.; Helquist, P.; Norrby, P.-O.; Wiest, O. *J. Chem. Theory Comput.* **2008**, *4*, 1313.

(5) Donoghue, P. J.; Helquist, P.; Norrby, P.-O.; Wiest, O. *J. Am. Chem. Soc.* **2008**, *131*, 410.

(6) Brandt, P.; Norrby, P.-O.; Andersson, P. G. *Tetrahedron* **2003**, *59*, 9695.

(7) Landis, C. R.; Halpern, J. *J. Am. Chem. Soc.* **1987**, *109*, 1746.

(8) Landis, C. R.; Steven, F. *Angew. Chem. Int. Ed.* **2000**, *39*, 2863.

(9) Watkins, A. L.; Landis, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 10306.

(10) Liu, P.; Yang, X.; Birman, V. B.; Houk, K. N. *Org. Lett.* **2012**, *14*, 3288.

(11) Liang, J.; Ruble, J. C.; Fu, G. C. *J. Org. Chem.* **1998**, *63*, 3154.

(12) Birman, V. B.; Li, X. *Org. Lett.* **2006**, *8*, 1351.

(13) Zuend, S. J.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2007**, *129*, 15872.

(14) Zuend, S. J.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2009**, *131*, 15358.

(15) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *The Journal of Chemical Physics* **2010**, *132*, 154104.

(16) Fischer, F. R.; Wood, P. A.; Allen, F. H.; Diederich, F. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17290.

(17) Choudhary, A.; Gandla, D.; Krow, G. R.; Raines, R. T. *J. Am. Chem. Soc.* **2009**, *131*, 7244.

(18) Allen, F. H.; Baalham, C. A.; Lommerse, J. P. M.; Raithby, P. R. *Acta Crystallographica Section B* **1998**, *54*, 320.

(19) Fristrup, P.; Tanner, D.; Norrby, P.-O. *Chirality* **2003**, *15*, 360.

(20) P.-O, N. *J. Mol. Struc. -Theochem* **2000**, *506*, 9.

(21) Nelson, D. W.; Gypser, A.; Ho, P. T.; Kolb, H. C.; Kondo, T.; Kwong, H.-L.; McGrath, D. V.; Rubin, A. E.; Norrby, P.-O.; Gable, K. P.; Sharpless, K. B. *J. Am. Chem. Soc.* **1997**, *119*, 1840.

(22) Norrby, P.-O.; Jensen, K. *Int. J. Pep. Res. Ther.* **2006**, *12*, 335.

(23) Norrby, P.-O.; Kolb, H. C.; Sharpless, K. B. *J. Am. Chem. Soc.* **1994**, *116*, 8470.

(24) Norrby, P.-O.; Rasmussen, T.; Haller, J.; Strassner, T.; Houk, K. N. *J. Am. Chem. Soc.* **1999**, *121*, 10186.

(25) Norrby, P.-O.; Wärnmark, K.; Åkermark, B.; Moberg, C. *J. Comput. Chem.* **1995**, *16*, 620.

(26) Kieken, E.; Wiest, O.; Helquist, P.; Cucciolito, M. E.; Flores, G.; Vitagliano, A.; Norrby, P.-O. *Organometallics* **2005**, *24*, 3737.

(27) Rasmussen, T.; Norrby, P.-O. *J. Am. Chem. Soc.* **2001**, *123*, 2464.

(28) Uyeda, C.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2011**, *133*, 5062.

(29) Brown, J. M.; Deeth, R. J. *Angew. Chem. Int. Ed.* **2009**, *48*, 4476.

(30) Nielsen, R. J.; Keith, J. M.; Stoltz, B. M.; Goddard, W. A. *J. Am. Chem. Soc.* **2004**, *126*, 7967.

(31) Huang, J.; Ianni, J. C.; Antoline, J. E.; Hsung, R. P.; Kozlowski, M. C. *Org. Lett.* **2006**, *8*, 1565.

(32) Lipkowitz, K. B.; Kozlowski, M. C. *Synlett* **2003**, *2003*, 1547.

(33) Lipkowitz, K. B.; D'Hue, C. A.; Sakamoto, T.; Stack, J. N. *J. Am. Chem. Soc.* **2002**, *124*, 14255.

(34) Lipkowitz, K. B.; Sakamoto, T.; Stack, J. *Chirality* **2003**, *15*, 759.

(35) Denmark, S. E.; Gould, N. D.; Wolf, L. M. *J. Org. Chem.* **2011**, *76*, 4337.

(36) Peña-Cabrera, E.; Norrby, P.-O.; Sjögren, M.; Vitagliano, A.; De Felice, V.; Oslob, J.; Ishii, S.; O'Neill, D.; Åkermark, B.; Helquist, P. *J. Am. Chem. Soc.* **1996**, *118*, 4299.

(37) Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.

(38) Hansch, C.; Leo, A. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, DC, 1995.

(39) Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*; University Science Books: Sausalito, 2006.

(40) Nic, M. J., J.; Kosata,B.; Jenkins A.; 2nd ed.; McNaught, A. D. W., A., Ed.; Blackwell Scientific Publications: Oxford, 2006.

(41) Seeman, J. I. *Chem. Rev.* **1983**, *83*, 83.

(42) Halpern, J. *Science (Washington, D. C., 1883-)* **1982**, *217*, 401.

(43) Zhang, W.; Loebach, J. L.; Wilson, S. R.; Jacobsen, E. N. *J. Am. Chem. Soc.* **1990**, *112*, 2801.

(44) Zhang, W.; Jacobsen, E. N. *J. Org. Chem.* **1991**, *56*, 2296.

(45) Jacobsen, E. N.; Zhang, W.; Muci, A. R.; Ecker, J. R.; Deng, L. *J. Am. Chem. Soc.* **1991**, *113*, 7063.

(46) Jacobsen, E. N.; Zhang, W.; Guler, M. L. *J. Am. Chem. Soc.* **1991**, *113*, 6703.

(47) Hansch, C.; Leo, A.; Taft, R. W. *Chem. Rev.* **1991**, *91*, 165.

(48) Palucki, M.; Finney, N. S.; Pospisil, P. J.; Güler, M. L.; Ishida, T.; Jacobsen, E. N. *J. Am. Chem. Soc.* **1998**, *120*, 948.

(49) Rodríguez-Escrich, S.; Reddy, K. S.; Jimeno, C.; Colet, G.; Rodríguez-Escrich, C.; Solà, L. s.; Vidal-Ferran, A.; Pericàs, M. A. *J. Org. Chem.* **2008**, *73*, 5340.

(50) Huang, Y.; Rawal, V. H. *Org. Lett.* **2000**, *2*, 3321.

(51) Huang, Y.; Rawal, V. H. *J. Am. Chem. Soc.* **2002**, *124*, 9662.

(52) Huang, Y.; Unni, A. K.; Thadani, A. N.; Rawal, V. H. *Nature* **2003**, *424*, 146.

(53) Thadani, A. N.; Stankovic, A. R.; Rawal, V. H. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5846.

(54) Unni, A. K.; Takenaka, N.; Yamamoto, H.; Rawal, V. H. *J. Am. Chem. Soc.* **2005**, *127*, 1336.

(55) Rajaram, S.; Sigman, M. S. *Org. Lett.* **2005**, *7*, 5473.

(56) Rajaram, S.; Sigman, M. S. *Org. Lett.* **2002**, *4*, 3399.

(57) Jensen, K. H.; Sigman, M. S. *Angew. Chem., Int. Ed.* **2007**, *46*, 4748.

(58) Jensen, K. H.; Sigman, M. S. *J. Org. Chem.* **2010**, *75*, 7194.

(59) Perrin, C. L. *Science* **1994**, *266*, 1665.

(60) Cleland, W.; Kreevoy, M. *Science* **1994**, *264*, 1887.

(61) Yoder, R. A.; Johnston, J. N. *Chem. Rev.* **2005**, *105*, 4730.

(62) Christianson, D. W. *Chem. Rev.* **2006**, *106*, 3412.

(63) Knowles, R. R.; Jacobsen, E. N. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 20678.

(64) Knowles, R. R.; Lin, S.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2010**, *132*, 5030.

(65) Vijay, D.; Sastry, G. N. *PCCP* **2008**, *10*, 582.

(66) Cubero, E.; Luque, F. J.; Orozco, M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5976.

(67) Calderone, C. T.; Williams, D. H. *J. Am. Chem. Soc.* **2001**, *123*, 6262.

(68) Wang, J.; Liu, X.; Feng, X. *Chem. Rev.* **2011**, *111*, 6947.

(69) Sigman, M. S.; Jacobsen, E. N. *J. Am. Chem. Soc.* **1998**, *120*, 4901.

(70) Vachal, P.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2002**, *124*, 10012.

(71) Vachal, P.; Jacobsen, E. N. *Org. Lett.* **2000**, *2*, 867.

(72) Zuend, S. J.; Coughlin, M. P.; Lalonde, M. P.; Jacobsen, E. N. *Nature* **2009**, *461*, 968.

(73) Miller, J. J.; Sigman, M. S. *Angew. Chem., Int. Ed.* **2008**, *47*, 771.

(74) Fürstner, A.; Shi, N. *J. Am. Chem. Soc.* **1996**, *118*, 12349.

(75) Miller, J. J.; Rajaram, S.; Pfaffenroth, C.; Sigman, M. S. *Tetrahedron* **2009**, *65*, 3110.

(76) Miller, J. J.; Sigman, M. S. *J. Am. Chem. Soc.* **2007**, *129*, 2752.

(77) Charton, M. *J. Org. Chem.* **1976**, *41*, 2217.

(78) Charton, M. *J. Am. Chem. Soc.* **1975**, *97*, 3691.

(79) Charton, M. *J. Am. Chem. Soc.* **1975**, *97*, 3694.

(80) Charton, M. *J. Am. Chem. Soc.* **1975**, *97*, 1552.

(81) Fujita, T.; Takayama, C.; Nakajima, M. *J. Org. Chem.* **1973**, *38*, 1623.

(82) *Steric Effects in Organic Chemistry*; Newman, M. S., Ed.; Wiley: New York, 1956.

(83) Taft, R. W., Jr. *J. Am. Chem. Soc.* **1953**, *75*, 4538.

(84) Sigman, M. S.; Miller, J. J. *J. Org. Chem.* **2009**, *74*, 7633.

(85) Giri, S.; Wang, D. Z.; Chattaraj, P. K. *Tetrahedron* **2010**, *66*, 4560.

(86) Li, X.; Deng, H.; Zhang, B.; Li, J.; Zhang, L.; Luo, S.; Cheng, J.-P. *Chem. -Eur. J.* **2010**, *16*, 450.

(87) Mikami, K.; Motoyama, Y.; Terada, M. *J. Am. Chem. Soc.* **1994**, *116*, 2812.

CHAPTER 2

DEVELOPMENT OF STERIC-BASED THREE DIMENSIONAL LINEAR FREE ENERGY RELATIONSHIPS IN NOZAKI-HIYAMA-KISHI ALLYLATION REACTIONS

Introduction

The ubiquitous implication of steric effects in asymmetric catalysis makes quantifying these effects a desirable goal.[1-4] The correlation found between the carbamoyl substituent and enantioselectivity shown in Figure 1.28 demonstrated that Charton values could be used to quantify steric elements in a system.[5] The breaks in the LFERs (Figure 1.29) led to several hypotheses, which included a proposed global shift in catalyst conformation, overcrowding of the most selective site on the chromium center funneling reactivity through less selective pathways, and flawed application of the Charton parameters. To fully explore these hypotheses, we embarked on a more sophisticated study of the ligand-catalyst system.

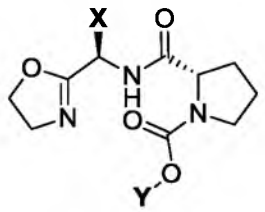
The study we designed was a simultaneous study of substituents X and Y shown in Figure 2.1. In prior studies, we had observed a subtle effect of position X on enantioselectivity. The nature of this effect was not intuitive because in the limited computation and hand modeling we had performed it appeared that substituent X was positioned away from the reactive center. This position had been dubbed the backbone position because of this presumed orientation. We thought probing the X position simultaneously with the Y position's more pronounced steric effects might provide information about how these two groups are interacting. Similarly, we thought by adding another dimension to our analysis we could

develop a model with greater predictive power than exhibited by the previous steric-based LFERs. With these goals, we set out to synthesize a ligand library where positions X and Y were systematically varied, evaluate these ligands in prototypical reactions, and collectively model the data as a complete set.[6]

Library Design and Synthesis

The initial ligand library was designed around the modular ligand scaffold described in our report of the Nozaki-Hiyama-Kishi allylation of ketones (Figure 2.1).[7] This truncated ligand scaffold was selected rather than the scaffold with which we had reported the steric-based LFERs (Figure 1.26).[5] This ligand scaffold presents two advantages over the previously reported scaffold. The first is that the truncated scaffold has only two diastereomers, one of which is not effective in the reaction in terms of enantioselectivity, which allows for more direct conclusions to be made about the nature of the steric effects. The second advantage is that the two steric effects we would be examining would be associated with the two chiral centers again potentially simplifying the analysis. The truncated ligand synthesis is compatible with an efficient route to the ligand library.[8] To assay the ligand library, we chose the allylations of benzaldehyde and acetophenone as model reactions. These reactions were selected primarily because of the wealth of data we had compiled over years of study that would allow us to benchmark the library approach. Also, the appropriate assays for enantioselectivity were already in place.

The substituents chosen for both the X and Y position were selected to reflect those commonly examined in an empirical examination of steric effects. Only carbon-based substituents were selected to avoid potential interaction between heteroatoms and the Cr-center, which might result in a mechanistic change. At the X position, the substituents depicted in Figure 2.2 were chosen primarily due to their availability from commercial amino acids. At



1

Figure 2.1. The oxazoline-proline modular ligand scaffold.

the Y position, the substituents shown in Figure 2.2 were selected to mirror those of the X position with exclusion of hydrogen because of the instability of such a molecule. The 4-heptyl (Hep) substituent was included because it was one of the substituents that "broke" from linearity in the previous LFERs. The underlying purpose for including this substituent was to generate some predictive power around this break in linearity and explore if the break was present across the ligand series.

Synthesis of the ligands began by installing the various carbamoyl substituents onto L-proline. Boc-L-proline is commercially available so synthesis of the Y_{tBu} ligand series forgoes the first step. The mild condensation of chloroformates onto L-proline is a straightforward way to generate the methyl-, ethyl-, and isopropyl-protected prolines (Scheme 2.1). The 4-heptyl chloroformate is not commercially available so an alternate path was pursued. To synthesize the desired proline analog (Hep), 4-heptanol is condensed with 4-nitrophenylchloroformate to prepare the carbonate **2** (Scheme 2.2). To avoid formation of the mixed anhydride, the benzyl ester-protected proline is used. The condensation of carbonate **2** with benzyl protected proline proceeds cleanly under the same conditions used with chloroformates. Finally, hydrogenolysis of the benzyl ester generates the acid, which can be carried on according to Scheme 2.1. Anderson's conditions for peptide bond formation gave good results for each of the differentially-protected prolines with the appropriate amino acid salts.[9,10] The final peptide bond can be formed under thermal conditions with aminoethanol in refluxing THF/toluene (1:1) mixture. Several reaction conditions were employed for oxazoline cyclization but conditions reported by Wipf and coworkers, using diethylaminosulphurtriflouride (DAST), were synthetically favorable.[11] The high cost of D-*tert*-leucine prompted us to synthesize the opposite enantiomer of the ligand for each of the X_{tBu} series. The same reactions outlined above were

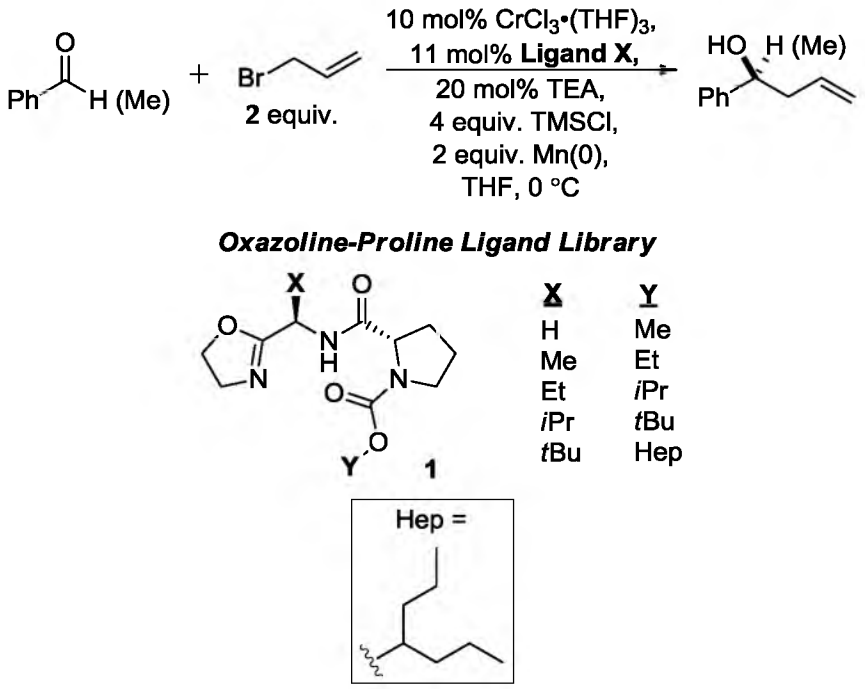
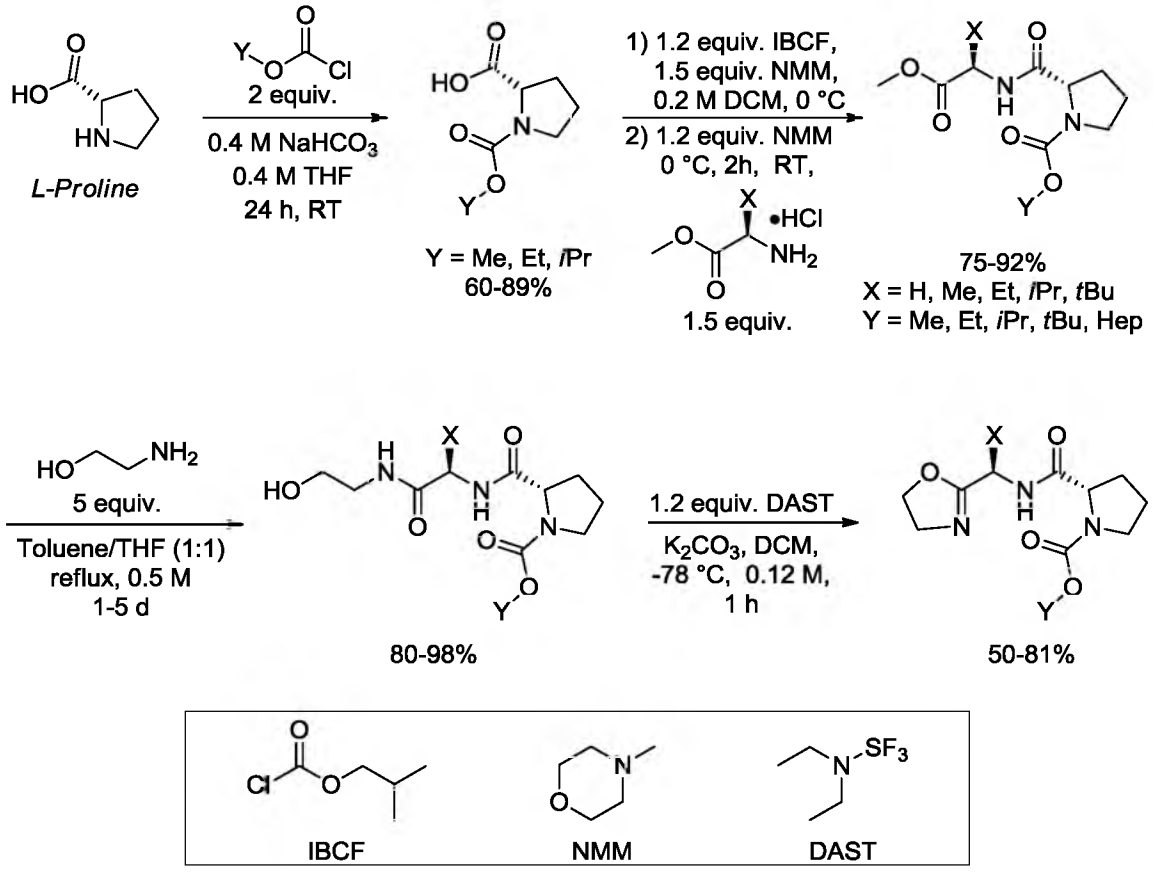
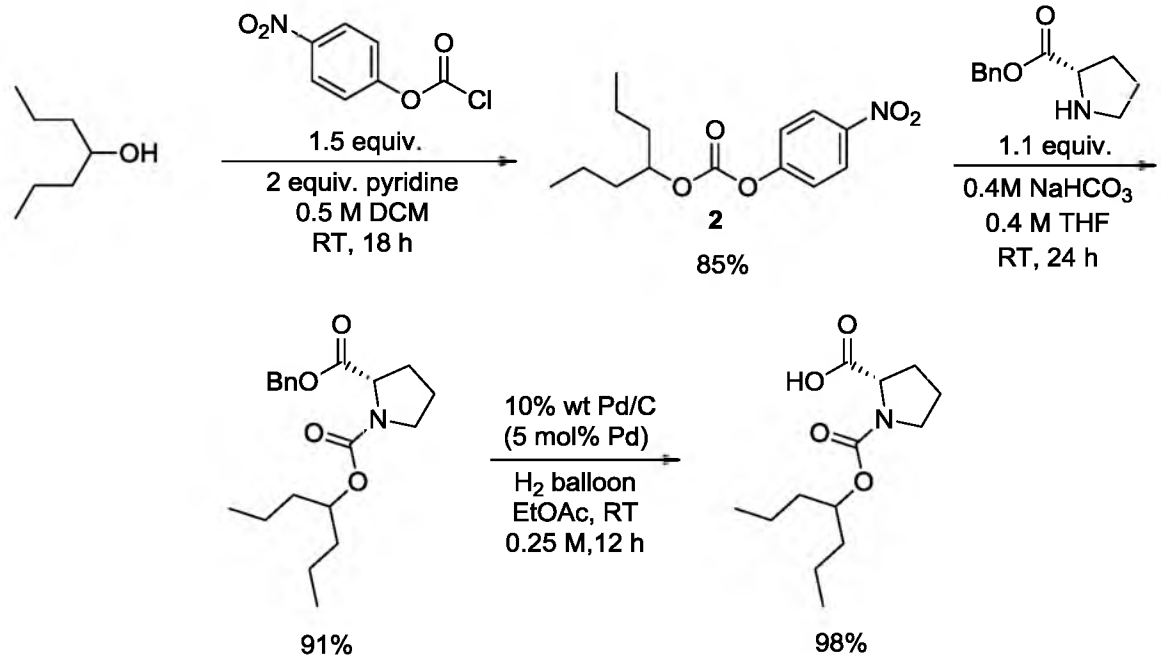


Figure 2.2. The allylation of benzaldehyde and acetophenone under standard conditions using the oxazoline-proline ligand library.

Scheme 2.1. Ligand library synthesis.



Scheme 2.2. Synthesis of the 4-heptyl substituted ligands.



performed with the appropriate D-proline derivatives. Therefore, the absolute value of er was used for the purposes described below.

With the 25 ligands synthesized, evaluation of the allylation of benzaldehyde was conducted under the conditions shown in Figure 2.2. Each allylation was repeated and the reported values represent the average of at least two replicates. The measured enantioselectivities are given in Table 2.1. To initially quantify the steric effects, Charton values were investigated.[12-15] This choice was based on our previous LFERs as well as our desire to extrapolate the model making use of the extensive library of reported Charton values to facilitate this extrapolation.[5,16] A scatterplot of the Charton values of X, Y and enantioselectivities is shown in Figure 2.3. Although our intention was to model the data collectively, we initially modeled the data in 2D slices in order to understand the overall character. Excitingly, a new LFER was discovered for the X_{Me} ligand series (Figure 2.4). Notably, this LFER included the $X_{Me}Y_{Hep}$ ligand in the correlation, a Y substituent, which had not been correlated in previous LFERs. This result suggested that extrapolation of this X_{Me} ligand series might not suffer from the same breaks in linearity as previously observed.

Developing a Model

As previously stated, our desire was to mathematically model the data as a complete set. In order to develop such a model, key assumptions required consideration. The first assumption is that linear free energy relationships refer to "linear" in the mathematical sense.[17] In order for a function to be linear in a mathematical sense it must possess two characteristics; the first is additivity and the second is homogeneity. Additivity can be defined such that the sum of functions is the same as each function added to the other as shown in Equation 2.1.

$$f(x+y) = f(x) + f(y) \quad (2.1)$$

Table 2.1. Results of the ligand library screen in the allylation of benzaldehyde.

Substituent		***er***	$\Delta\Delta G^{\ddagger}$
X_{H}	Y_{Me}	50:50	0.00
X_{H}	Y_{Et}	36.5:63.5	0.30
X_{H}	Y_{iPr}	38.5:61.5	0.25
X_{H}	Y_{tBu}	15:85	0.94
X_{H}	Y_{Hep}	26.3:73.7	0.56
X_{Me}	Y_{Me}	44:56	0.13
X_{Me}	Y_{Et}	39:61	0.24
X_{Me}	Y_{iPr}	35.5:64.5	0.32
X_{Me}	Y_{tBu}	17:83	0.86
X_{Me}	Y_{Hep}	12:88	1.08
X_{Et}	Y_{Me}	39:61	0.24
X_{Et}	Y_{Et}	25.5:74.5	0.58
X_{Et}	Y_{iPr}	16.5:83.5	0.88
X_{Et}	Y_{tBu}	18:82	0.82
X_{Et}	Y_{Hep}	11:89	1.13
X_{iPr}	Y_{Me}	37.5:62.5	0.28
X_{iPr}	Y_{Et}	16.5:83.5	0.88
X_{iPr}	Y_{iPr}	14.5:85.5	0.96
X_{iPr}	Y_{tBu}	8.5:91.5	1.29
X_{iPr}	Y_{Hep}	24.3:75.7	0.62
X_{tBu}	Y_{Me}	39.6:60.4	0.23
X_{tBu}	Y_{Et}	41:59	0.20
X_{tBu}	Y_{iPr}	35.3:64.7	0.33
X_{tBu}	Y_{tBu}	35:65	0.34
X_{tBu}	Y_{Hep}	31.5:68.5	0.42

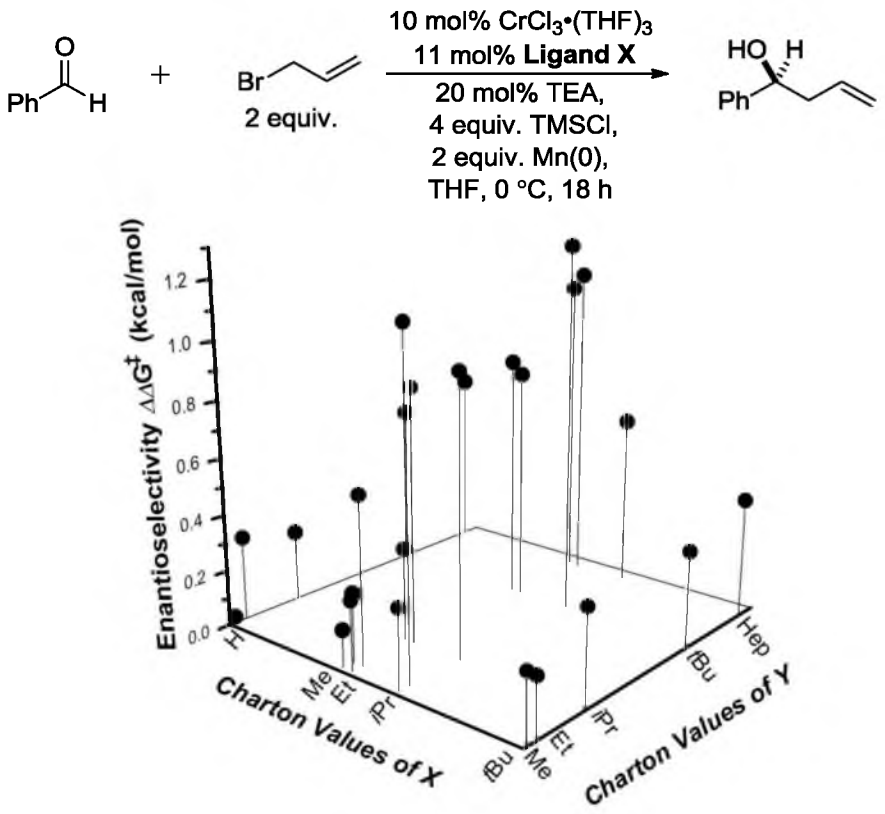


Figure 2.3. Three-dimensional scatterplot of the measured enantioselectivities in the allylation of benzaldehyde.

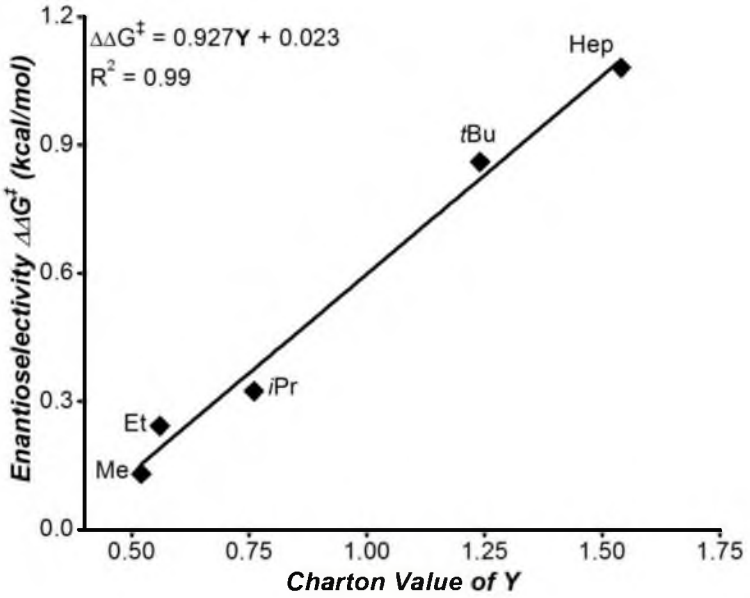


Figure 2.4. LFER between the Y substituent and enantioselectivity which includes the 4-Heptyl (Hep) derived ligand.

Homogeneity implies that for each discreet input into the function there is a single output as demonstrated in Equation 2.2.

$$f(\alpha x) = \alpha f(x) \quad (2.2)$$

Simply, a linear function is any function where one value in X corresponds to one value in Y. Using LFERs in asymmetric catalysis assumes homogeneity and additivity for the mechanism of asymmetric induction for the reaction but also the source of the parameters used to develop the LFER. These assumptions must be made with caution if causality is going to be implied. Equation 2.3, upon which all LFERs are based, is a linear function.

$$\Delta\Delta G^{\ddagger} = -RTln(k_{rel}) \quad (2.3)$$

A linear free energy relationship does not have to exhibit a first order (straight line) relationship, i.e. polynomials and other mathematical operations can produce a linear function. First order relationships have prevailed in the literature because interpretation and development is simplified. For this reason, the term "linear" is often interpreted to mean only first order character. For the remainder of this thesis, the term "linear" will reference the mathematical definition and not the colloquial interpretation.

Justification for including not first order terms is found in examining the Lennard-Jones potential, a model for steric repulsion. Steric effects, by definition, are the electrostatic repulsive forces exhibited by two atoms or molecules as they approach one another beyond their Van der Waals distances, as modeled by the Lennard-Jones potential (Figure 2.5).[18-20] Thus, any comprehensively modeled steric interaction would not likely possess straight line character but rather a curved exponential character. However, this does not discount the first order relationships previously described. The complexity of steric interactions can be modeled in many different ways, but should not be limited to first order relationships with steric parameters (Figure 1.27).

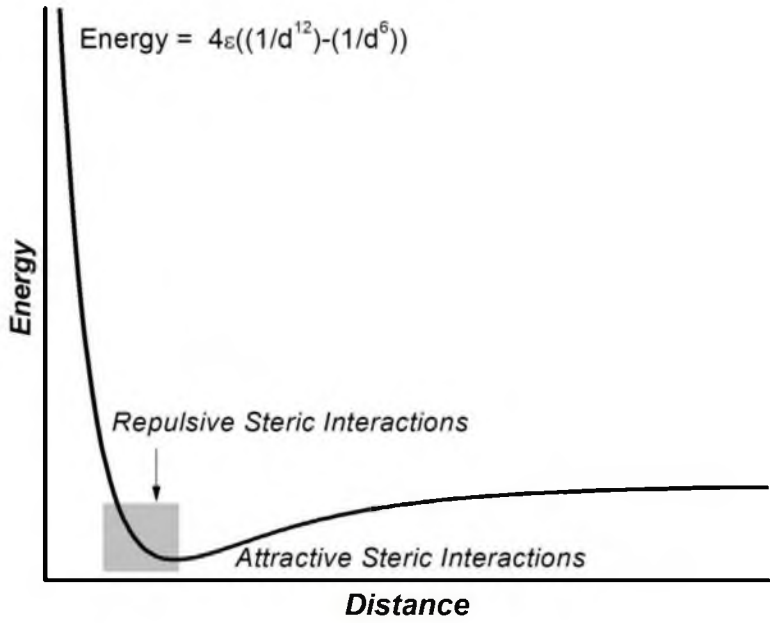


Figure 2.5. The Lennard-Jones potential for steric interactions indicating the region where straight line character might be approximated for a repulsive interaction.

Our empirical inspection of the data in Figure 2.3 led us to propose a full third order polynomial as the base model, upon which we could regress the data (Equation 2.4).[21]

$$\Delta\Delta G^{\ddagger} = z0 + aX + bY + cX^2 + dY^2 + fXY + gX^3 + hY^3 + iYX^2 + jXY^2 \quad (2.4)$$

The data from the ligand library would possess 25 individual points gathered from the 25 individual ligands. A base model with higher order terms would mean a larger number of terms to fit to the data and could lead to over fitting the model to the data. The number of data points less the number of terms used to model the data is defines the number of degrees of freedom. The higher the degrees of freedom, the more ability a model possesses to accommodate experimental error. Thus, a balance between the terms included and the size of the data set must be achieved. The fear of over fitting a model and losing predictive power led to limiting the base model as a full third order polynomial. We chose a full 3rd order polynomial to balance an apparent need of higher order terms to model the curvature evident in the data, and maximizing the degrees of freedom.

Equation 2.4 includes crossterms which relate variables X and Y under a single coefficient. Crossterms of this variety are typically included in models where the variables are presumed to be interacting. Hence, crossterms were included into the model as a test of the hypothesis that the two steric elements were synergistically influencing enantioselectivity. This hypothesis would be tested through regression of the data onto Equation 2.4; if the crossterms were deemed statistically significant to a predictive model, then the hypothesis would be supported.

To fit the polynomial model to the data, a backward stepwise regression technique was used. Applying regression techniques to polynomial models requires that the Charton values (υ) be translated to center around zero. This translation is required to eliminate a potential source of error. When only positive values are used, second and third order polynomial character is

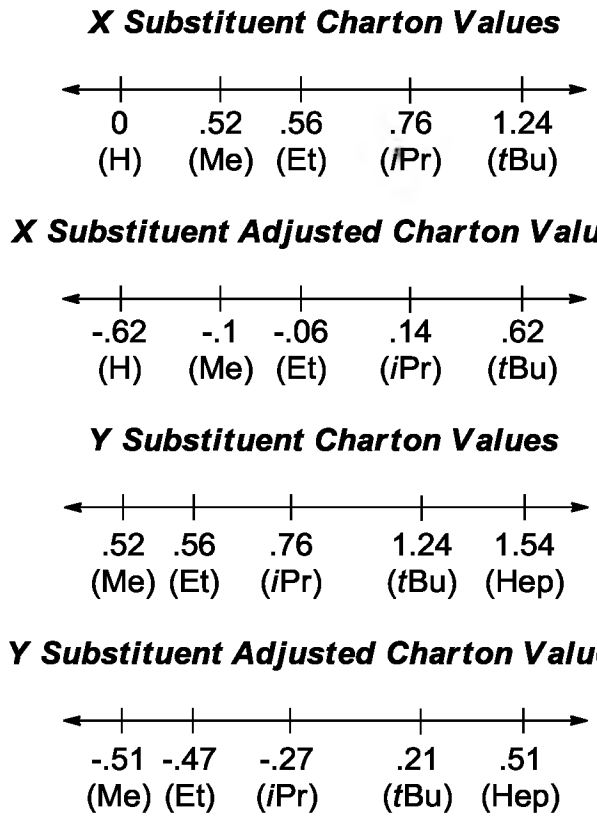
difficult to distinguish using standard regression techniques and the true character of the relationship can be obscured. The translation was accomplished according to Equation 2.5 for values in each dimension and Figure 2.6 lists the adjusted values.

$$\upsilon_{adjusted} = \upsilon_{original} - (\upsilon_{max} - \upsilon_{min})/2 \quad (2.5)$$

Because Charton values are relative values, which were arbitrarily set with hydrogen equal to zero, the translation has no effect on their interpretation.

To perform the regression, the data was assembled into a design matrix X and a response matrix Y shown in Figure 2.7.[21] The columns of the design matrix correspond to the terms within a 3rd order polynomial base model, where the first column is an offset (z0) and subsequent columns correspond to X, Y, X^2, Y^2, XY, etc. In this case, the response matrix is only a single column of the measured enantioselectivity converted to $\Delta\Delta G^{\ddagger}$ values. The rows in the design matrix correspond to each different ligand with adjusted Charton values for the X and Y substituent inserted into the columns in the appropriate order. For an example, for the ligand $X_{Me}Y_{Me}$, the adjusted value for Me along the x axis is -0.1 which is placed in the X column, and that value squared (0.01) is placed in the X^2 column. Similarly, the adjusted Charton value for Me along the Y axis is -0.51, so that value is placed in the Y column and similarly that value squared (0.26) is placed in the Y^2 column. A unique feature of translating the data along two separate axes is that two different values for each group emerge. Because the two data sets were designed with different substituents, the separate values are required. However, the key principle is that the values remain relative along a single axis in order to maintain comparability and not bias the model's development.

Examining the data in matrices can be instructive as to how multivariate regression is performed. The base 3rd order polynomial has ten unknown parameter coefficients (z0-j). These coefficients can be considered in the same manner that slope is considered in a first order



Substituent		Adj. Charton Values X	Y
X_{H}	Y_{Me}	-0.62	-0.51
X_{H}	Y_{Et}	-0.62	-0.47
X_{H}	Y_{iPr}	-0.62	-0.27
X_{H}	Y_{tBu}	-0.62	0.21
X_{H}	Y_{Hep}	-0.62	0.51
X_{Me}	Y_{Me}	-0.1	-0.51
X_{Me}	Y_{Et}	-0.1	-0.47
X_{Me}	Y_{iPr}	-0.1	-0.27
X_{Me}	Y_{tBu}	-0.1	0.21
X_{Me}	Y_{Hep}	-0.1	0.51
X_{Et}	Y_{Me}	-0.06	-0.51
X_{Et}	Y_{Et}	-0.06	-0.47
X_{Et}	Y_{iPr}	-0.06	-0.27
X_{Et}	Y_{tBu}	-0.06	0.21
X_{Et}	Y_{Hep}	-0.06	0.51
X_{iPr}	Y_{Me}	0.14	-0.51
X_{iPr}	Y_{Et}	0.14	-0.47
X_{iPr}	Y_{iPr}	0.14	-0.27
X_{iPr}	Y_{tBu}	0.14	0.21
X_{iPr}	Y_{Hep}	0.14	0.51
X_{tBu}	Y_{Me}	0.62	-0.51
X_{tBu}	Y_{Et}	0.62	-0.47
X_{tBu}	Y_{iPr}	0.62	-0.27
X_{tBu}	Y_{tBu}	0.62	0.21
X_{tBu}	Y_{Hep}	0.62	0.51

Figure 2.6. A) Original Charton values and their translation according to Equation 5 along both axes. B) Tabulation of translated values for the ligand library.

Ligand	Design Matrix X										Response Matrix Y
	z0	X	Y	X^2	Y^2	XY	X^3	X^3	YX^2	XY^2	$\Delta\Delta G^{\ddagger}$
X_H Y_{Me}	1	-0.62	-0.51	0.3844	0.2601	0.3162	-0.2383	-0.1327	-0.1960	-0.1613	0.00
X_H Y_{Et}	1	-0.62	-0.47	0.3844	0.2209	0.2914	-0.2383	-0.1038	-0.1807	-0.1370	0.30
X_H Y_{iPr}	1	-0.62	-0.27	0.3844	0.0729	0.1674	-0.2383	-0.0197	-0.1038	-0.0452	0.25
X_H Y_{tBu}	1	-0.62	0.21	0.3844	0.0441	-0.1302	-0.2383	0.0093	0.0807	-0.0273	0.94
X_H Y_{Hep}	1	-0.62	0.51	0.3844	0.2601	-0.3162	-0.2383	0.1327	0.1960	-0.1613	0.56
X_{Me} Y_{Me}	1	-0.1	-0.51	0.01	0.2601	0.051	-0.0010	-0.1327	-0.0051	-0.0260	0.13
X_{Me} Y_{Et}	1	-0.1	-0.47	0.01	0.2209	0.047	-0.0010	-0.1038	-0.0047	-0.0221	0.24
X_{Me} Y_{iPr}	1	-0.1	-0.27	0.01	0.0729	0.027	-0.0010	-0.0197	-0.0027	-0.0073	0.32
X_{Me} Y_{tBu}	1	-0.1	0.21	0.01	0.0441	-0.021	-0.0010	0.0093	0.0021	-0.0044	0.86
X_{Me} Y_{Hep}	1	-0.1	0.51	0.01	0.2601	-0.051	-0.0010	0.1327	0.0051	-0.0260	1.08
X_{Et} Y_{Me}	1	-0.06	-0.51	0.0036	0.2601	0.0306	-0.0002	-0.1327	-0.0018	-0.0156	0.24
X_{Et} Y_{Et}	1	-0.06	-0.47	0.0036	0.2209	0.0282	-0.0002	-0.1038	-0.0017	-0.0133	0.58
X_{Et} Y_{iPr}	1	-0.06	-0.27	0.0036	0.0729	0.0162	-0.0002	-0.0197	-0.0010	-0.0044	0.88
X_{Et} Y_{tBu}	1	-0.06	0.21	0.0036	0.0441	-0.0126	-0.0002	0.0093	0.0008	-0.0026	0.82
X_{Et} Y_{Hep}	1	-0.06	0.51	0.0036	0.2601	-0.0306	-0.0002	0.1327	0.0018	-0.0156	1.13
X_{iPr} Y_{Me}	1	0.14	-0.51	0.0196	0.2601	-0.0714	0.0027	-0.1327	-0.0100	0.0364	0.28
X_{iPr} Y_{Et}	1	0.14	-0.47	0.0196	0.2209	-0.0658	0.0027	-0.1038	-0.0092	0.0309	0.88
X_{iPr} Y_{iPr}	1	0.14	-0.27	0.0196	0.0729	-0.0378	0.0027	-0.0197	-0.0053	0.0102	0.96
X_{iPr} Y_{tBu}	1	0.14	0.21	0.0196	0.0441	0.0294	0.0027	0.0093	0.0041	0.0062	1.29
X_{iPr} Y_{Hep}	1	0.14	0.51	0.0196	0.2601	0.0714	0.0027	0.1327	0.0100	0.0364	0.62
X_{tBu} Y_{Me}	1	0.62	-0.51	0.3844	0.2601	-0.3162	0.2383	-0.1327	-0.1960	0.1613	0.23
X_{tBu} Y_{Et}	1	0.62	-0.47	0.3844	0.2209	-0.2914	0.2383	-0.1038	-0.1807	0.1370	0.20
X_{tBu} Y_{iPr}	1	0.62	-0.27	0.3844	0.0729	-0.1674	0.2383	-0.0197	-0.1038	0.0452	0.33
X_{tBu} Y_{tBu}	1	0.62	0.21	0.3844	0.0441	0.1302	0.2383	0.0093	0.0807	0.0273	0.34
X_{tBu} Y_{Hep}	1	0.62	0.51	0.3844	0.2601	0.3162	0.2383	0.1327	0.1960	0.1613	0.42

Figure 2.7. The design matrix used in the initial iteration of model development for the allylation of benzaldehyde.

model, the value of the coefficient reflects the sensitivity of the model to the term. The data can be considered a series of equations (rows) with solutions ($\Delta\Delta G^{\ddagger}$ values) that can be used to solve for the ten unknown parameter coefficients. As long as there are fewer unknowns than equations, linear least squares regression can be performed. Accordingly, regression was performed using the linear algebra definition of regression given in Equation 2.6, where C is the output matrix of coefficients *z0-j.*

$$C = (X^T X)^{-1}(X^T Y) \quad (2.6)$$

To analyze the regression model, the variance-covariance matrix V was also calculated according to Equation 2.7 where s^2 is the variance of the $\Delta\Delta G^{\ddagger}$ values.

$$V = s^2(X^T X)^{-1} \quad (2.7)$$

The variance-covariance matrix estimates the variance or error associated with each coefficient value along the diagonal. The off diagonal terms relate the covariance between terms in the model. Covariance between terms implies two terms depend on each other linearly. Linear dependence between terms is a major source of error in predictive models because it obscures the true relationships within a model. Ideally, no covariance between terms would lead to the most predictive model. Hence, interpretation of the variance-covariance matrix was an essential component of model development.

To demonstrate how the model was developed the first iteration will be given in detail, and the remaining iterations can be examined in the supplemental information. The original data is given in Figure 2.7 as matrices X and Y. Regression according to Equation 2.6 was performed and the resulting coefficient estimates are given in Equation 2.8.

$$\Delta\Delta G^{\ddagger} = 0.93 - 0.0001X + 0.58Y - 0.91X^2 - 1.01Y^2 - 0.50XY - 0.41X^3 - 0.0002Y^3 - 0.48YX^2 + 0.00XY^2$$

(2.8)

This model was used to predict $\Delta\Delta G^{\ddagger}$ values for all 25 data points, which were plotted against the experimentally measured $\Delta\Delta G^{\ddagger}$ values as shown in Figure 2.8. The slope of the linear correlation between predicted and measured values is the R^2 value for the model and is 0.75. This value was used as a measure for the statistical goodness of fit. To avoid over fitting the data, an analysis of variance (ANOVA) was performed and the Fischer statistic (f-value) for the model was calculated to be 46. These two statistical measures provide the criteria, we used to development the initial model. In order to maximize each of these statistical criteria the variance-covariance matrix was calculated and used as rational for elimination of terms (Figure 2.9). Examining the highlighted values along the diagonal reveals that jXY^2, which has a value of zero, also has the highest error relative to the coefficient value. As a result, the column corresponding to this term in the design matrix X is removed giving a revised design matrix X_1 shown in Figure 2.10. This design matrix is used to evaluate new coefficient terms given in Equation 2.9.

$$\Delta\Delta G^{\ddagger} = 0.93 - 0.0001X + 0.58Y - 0.91X^2 - 1.01Y^2 - 0.50XY - 0.41X^3 - 0.0002Y^3 - 0.48YX^2 \quad (2.9)$$

Elimination of this zero term did not have any significant effect on the terms of the model but does increase the f-value to 58. The new variance-covariance matrix shown in Figure 2.11 reveals the term with the highest relative error is the hY^3 term, which would be eliminated in subsequent rounds. This process was repeated until the elimination of a term led to an insignificant or detrimental effect on the statistical criteria R^2 and f-value. It should also be noted that in cases where omission of a term led to a detrimental effect on statistical criteria, elimination of the terms that showed covariance with the initial term was also performed. Once a minimum number of terms were reached, each individual term was added again to the model one by one and their statistical significance reassessed to ensure no term was eliminated due to high covariance unnecessarily. The final derived model is given in Equation 2.10. Equation 2.10

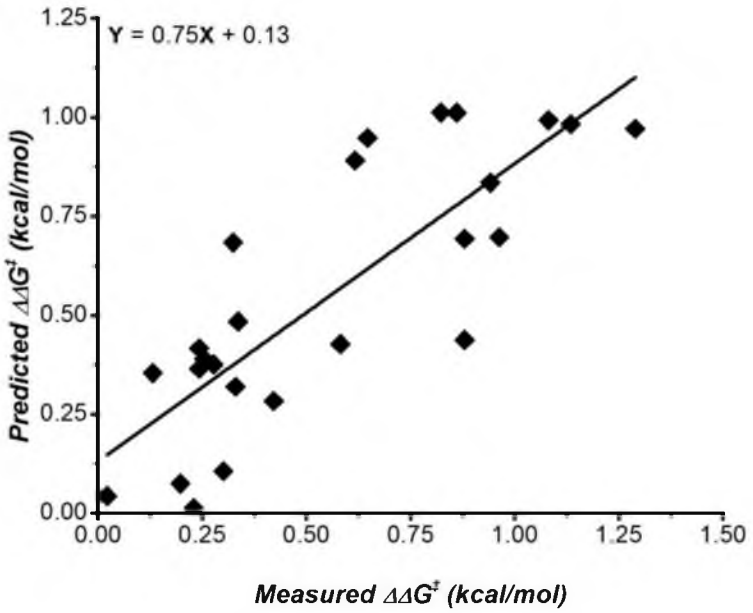


Figure 2.8. Plot comparing the predicted enantioselectivities from Equation 8 to experimentally measured enantioselectivities.

z0	X	Y	X^2	Y^2	XY	X^3	X^3	YX^2	XY^2
0.21	0.07	0.00	-0.21	-0.78	0.00	-0.17	0.13	-0.13	-0.02
0.07	7.41	0.00	-0.21	-0.02	0.02	-17.13	0.00	0.00	-4.84
0.00	0.00	3.13	-0.13	0.37	0.01	0.00	-11.96	-1.20	0.00
-0.21	-0.21	-0.13	1.29	0.00	0.00	0.54	0.00	0.79	0.00
-0.78	-0.02	0.37	0.00	4.63	0.00	0.00	-1.04	0.00	0.12
0.00	0.02	0.01	0.00	0.00	1.58	0.00	0.00	-0.04	0.83
-0.17	-17.13	0.00	0.54	0.00	0.00	46.40	0.00	0.00	0.00
0.13	0.00	-11.96	0.00	-1.04	0.00	0.00	53.33	0.00	0.01
-0.13	0.00	-1.20	0.79	0.00	-0.04	0.00	0.00	7.45	0.00
-0.02	-4.84	0.00	0.00	0.12	0.83	0.00	0.01	0.00	28.70

Figure 2.9. The variance-covariance matrix associated with coefficients of Equation 8.

Ligand		Revised Design Matrix X_1									Response Matrix Y
		z0	X	Y	X^2	Y^2	XY	X^3	Y^3	YX^2	$\Delta\Delta G^{\ddagger}$
X_H	Y_{Me}	1	-0.62	-0.51	0.3844	0.2601	0.3162	-0.2383	-0.1327	-0.1960	0.00
X_H	Y_{Et}	1	-0.62	-0.47	0.3844	0.2209	0.2914	-0.2383	-0.1038	-0.1807	0.30
X_H	Y_{iPr}	1	-0.62	-0.27	0.3844	0.0729	0.1674	-0.2383	-0.0197	-0.1038	0.25
X_H	Y_{tBu}	1	-0.62	0.21	0.3844	0.0441	-0.1302	-0.2383	0.0093	0.0807	0.94
X_H	Y_{Hep}	1	-0.62	0.51	0.3844	0.2601	-0.3162	-0.2383	0.1327	0.1960	0.56
X_{Me}	Y_{Me}	1	-0.1	-0.51	0.01	0.2601	0.051	-0.0010	-0.1327	-0.0051	0.13
X_{Me}	Y_{Et}	1	-0.1	-0.47	0.01	0.2209	0.047	-0.0010	-0.1038	-0.0047	0.24
X_{Me}	Y_{iPr}	1	-0.1	-0.27	0.01	0.0729	0.027	-0.0010	-0.0197	-0.0027	0.32
X_{Me}	Y_{tBu}	1	-0.1	0.21	0.01	0.0441	-0.021	-0.0010	0.0093	0.0021	0.86
X_{Me}	Y_{Hep}	1	-0.1	0.51	0.01	0.2601	-0.051	-0.0010	0.1327	0.0051	1.08
X_{Et}	Y_{Me}	1	-0.06	-0.51	0.0036	0.2601	0.0306	-0.0002	-0.1327	-0.0018	0.24
X_{Et}	Y_{Et}	1	-0.06	-0.47	0.0036	0.2209	0.0282	-0.0002	-0.1038	-0.0017	0.58
X_{Et}	Y_{iPr}	1	-0.06	-0.27	0.0036	0.0729	0.0162	-0.0002	-0.0197	-0.0010	0.88
X_{Et}	Y_{tBu}	1	-0.06	0.21	0.0036	0.0441	-0.0126	-0.0002	0.0093	0.0008	0.82
X_{Et}	Y_{Hep}	1	-0.06	0.51	0.0036	0.2601	-0.0306	-0.0002	0.1327	0.0018	1.13
X_{iPr}	Y_{Me}	1	0.14	-0.51	0.0196	0.2601	-0.0714	0.0027	-0.1327	-0.0100	0.28
X_{iPr}	Y_{Et}	1	0.14	-0.47	0.0196	0.2209	-0.0658	0.0027	-0.1038	-0.0092	0.88
X_{iPr}	Y_{iPr}	1	0.14	-0.27	0.0196	0.0729	-0.0378	0.0027	-0.0197	-0.0053	0.96
X_{iPr}	Y_{tBu}	1	0.14	0.21	0.0196	0.0441	0.0294	0.0027	0.0093	0.0041	1.29
X_{iPr}	Y_{Hep}	1	0.14	0.51	0.0196	0.2601	0.0714	0.0027	0.1327	0.0100	0.62
X_{tBu}	Y_{Me}	1	0.62	-0.51	0.3844	0.2601	-0.3162	0.2383	-0.1327	-0.1960	0.23
X_{tBu}	Y_{Et}	1	0.62	-0.47	0.3844	0.2209	-0.2914	0.2383	-0.1038	-0.1807	0.20
X_{tBu}	Y_{iPr}	1	0.62	-0.27	0.3844	0.0729	-0.1674	0.2383	-0.0197	-0.1038	0.33
X_{tBu}	Y_{tBu}	1	0.62	0.21	0.3844	0.0441	0.1302	0.2383	0.0093	0.0807	0.34
X_{tBu}	Y_{Hep}	1	0.62	0.51	0.3844	0.2601	0.3162	0.2383	0.1327	0.1960	0.42

Figure 2.10. The first iteration of the design matrix after elimination of the hXY^2 term.

z0	X	Y	X^2	Y^2	XY	X^3	Y^3	YX^2
0.21	0.06	0.00	-0.21	-0.78	0.00	-0.17	0.13	-0.13
0.06	6.59	0.00	-0.21	0.00	0.16	-17.13	0.00	0.00
0.00	0.00	3.13	-0.13	0.37	0.01	0.00	-11.96	-1.20
-0.21	-0.21	-0.13	1.29	0.00	0.00	0.54	0.00	0.79
-0.78	0.00	0.37	0.00	4.63	0.00	0.00	-1.04	0.00
0.00	0.16	0.01	0.00	0.00	1.56	0.00	0.00	-0.04
-0.17	-17.13	0.00	0.54	0.00	0.00	46.40	0.00	0.00
0.13	0.00	-11.96	0.00	-1.04	0.00	0.00	53.33	0.00
-0.13	0.00	-1.20	0.79	0.00	-0.04	0.00	0.00	7.45

Figure 2.11. The variance-covariance matrix associated with the coefficients in Equation 9.

represents the achievement of our goal and models the entire data set. The final statistical measures for the model are an R^2 value of 0.75 and an f-value of 125. Equation 2.10 possesses three dimensions (adjusted Charton value of X, adjusted Charton value of Y, and experimentally determined $\Delta\Delta G^{\ddagger}$) so it can be visualized as a surface as shown in Figure 2.12.

$$\Delta\Delta G^{\ddagger} = 0.93 + 0.58Y - 0.91X^2 - 1.01Y^2 - 0.50XY - 0.41X^3 - 0.48YX^2 \quad (2.10)$$

The development of the model was based on statistical significance of the terms incorporated into the model. The result of this treatment of the data is that each term's has statistical significance in the final model. The inclusion of two statistically significant crossterms in the final model suggests that X and Y substituents are in fact interacting and influencing enantioselectivity. The spectroscopically elusive nature of the ligand-catalyst complex has deflected our attempts to probe structure directly but hand models and limited computation have suggested that the groups are on opposite sides of the catalyst. While the exact meaning of this relationship is still unclear, the crossterms have provided considerable insight into the catalyst system.

The purpose of the 3D free energy relationship was not only to examine the proposed synergistic effect but to overcome the observed breaks in linearity. The 3D model shown in Figure 2.12, and the 2D LFER shown in Figure 2.4 presented a situation to compare the predictive power of the two models. The two models exhibit different trends with increasing size in the Y dimension. The 3D model demonstrates downward curvature and the 2D model exhibits a positive trend. Our desire to extrapolate the two models as well as externally validate the 3D model led us to synthesize ligands **3a**-**3c** because of their large Y value would be a maximum extrapolation of the linear model. Ligand **3b** provides direct comparison between models and ligands **3a** and **3c** were also used to validate Equation 2.10. The synthesis paralleled the previously described syntheses and the results for these ligands are given in Figure 2.13. For

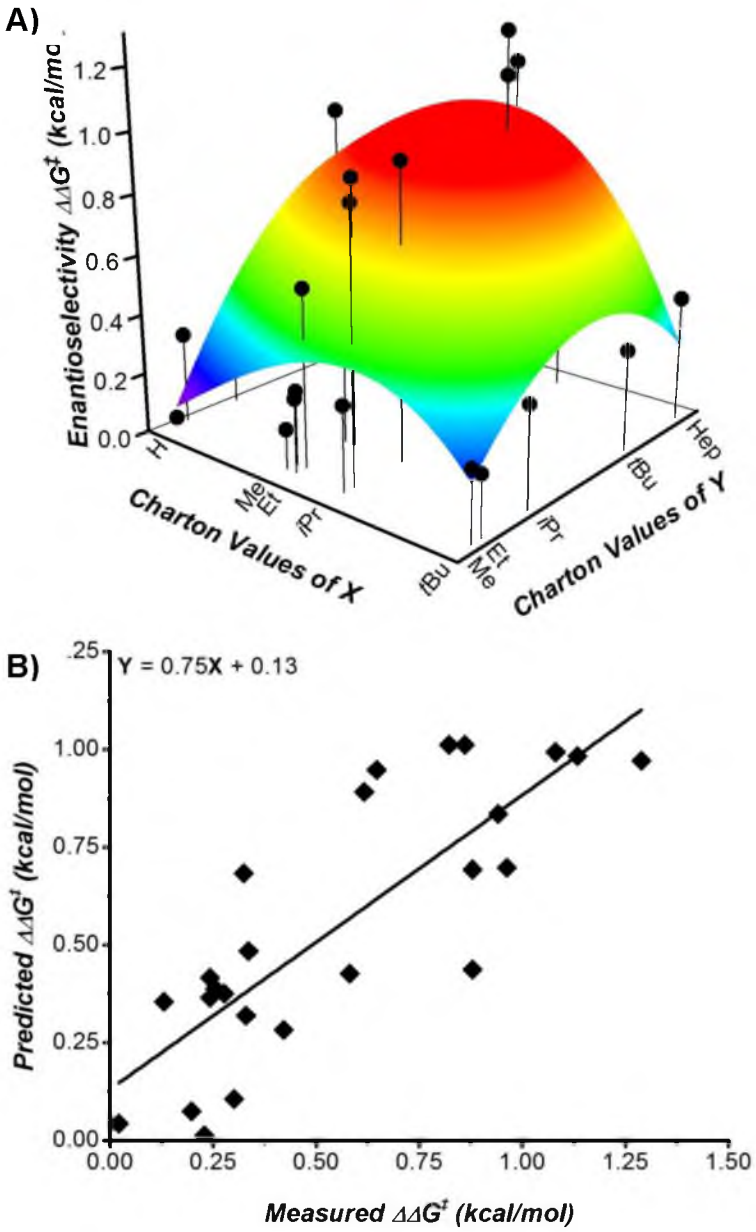
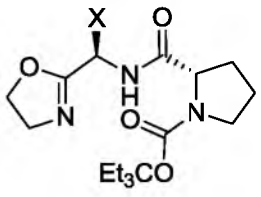


Figure 2.12. A) Three-dimensional surface model described by Equation 10 overlaid onto the original data. B) Plot of the predicted enantioselectivities given by Equation 10 with their experimentally measured values.



		Predicted er			*Measured*
	X	Linear	3D	*Error*	er
3a	H	NA	40.8:59.2	±4	42.5:57.5
3b	Me	3:97	46.7:53.3	±4	36:64
3c	*i*Pr	NA	38.6:61.3	±4	40:60

Figure 2.13. Comparison of predictions made by the 2D LFER and the 3D LFER for a several ligands bearing a bulky carbamoyl group.

ligand **3b**, the 3D LFER was able to predict the enantioselectivity with much greater accuracy than the 2D LFER. The 2D LFER relationship exhibits another break in linearity for a large substituent. Prediction of enantioselectivity for ligands **3a** and **3c**, is accurate to within error and prediction of ligand **3b** is within reason. The accuracy of these predictions is remarkable given the magnitude of the extrapolation and the limitations imposed on the model discussed below.

In critique of this effort to develop 3D LFERs, development of the library and data evaluation was time consuming and led to only 3 predictions of poor ligand performance. The model itself did not give a high degree of correlation to the source data. The method would be more compelling if it were capable of predicting higher performing ligands, whereas from a practical stand point most chemists would not invest this much effort for such predictions. The analysis did reveal the presence of crossterms relating X and Y. The use of crossterms in models such as this is an intriguing technique, which might be capable of probing relationships for which there is no direct measure. Similar to the examination of benzaldehyde, a separate evaluation was performed on data gathered from the allylation of acetophenone but weaker correlation and no predictive power was exhibited by several derived models. In order to refine the 3D-LFER approach, it became apparent that the principles of experimental design would need to be applied to the system.

The Principles of Experimental Design Applied to Asymmetric Catalysis

Experimental design has existed since the 1870s and has become a field unto itself.[21] The application of statistical inference in chemical problems has taken the name chemometrics or chemoinformatics. The fundamental concept is that if statistical inferences are going to be made about a body of data with variance, then statistical design of the experiment will lead to

greater predictive power. In order to develop better 3D-LFERs, two key principles were applied from the principles of experimental design namely the even distribution of data and interpolation of result rather than extrapolation.[21]

Due to some inherent complications, the principles of experimental design cannot be applied to steric effects and asymmetric catalysis in their purest forms. The conflict in this application primarily arises from the discreet nature of the substituents employed. Charton's steric parameters do not present a continuous spectrum of values but rather a series of discreet values limited by synthetic constraints. Essentially, there is no substituent between a methyl group and an ethyl group. While not quite correct, this statement exemplifies the problem inherent in applying experimental design principles to steric parameters. Experimental design in its purest form is applied to variables that are continuous such as temperature or pressure. This limitation not only complicates the application of the principles but also the interpretation of the results.

The first principle of experimental design applied to the ligand library dealt with the distribution of data points along the X and Y axes. Figure 2.14 shows how the data points were aligned according to our initial attempt. The data is not evenly spread across either axis. This uneven distribution results in biasing the model with greater predictive power in quadrant III and weaker predictive power in quadrant I. The bias in the model can complicate and cause inaccuracy in extrapolation. Experimental design dictates that a data set designed with an even spread of data points can have as much or greater predictive power than one with more points that is unevenly spread. One of the critiques of the previous model was that 25 (5×5) ligands required a large synthetic effort to generate moderate predictions. Revising the ligand library to include only 9 (3×3) ligands, which have substituents that are evenly spaced along both the X

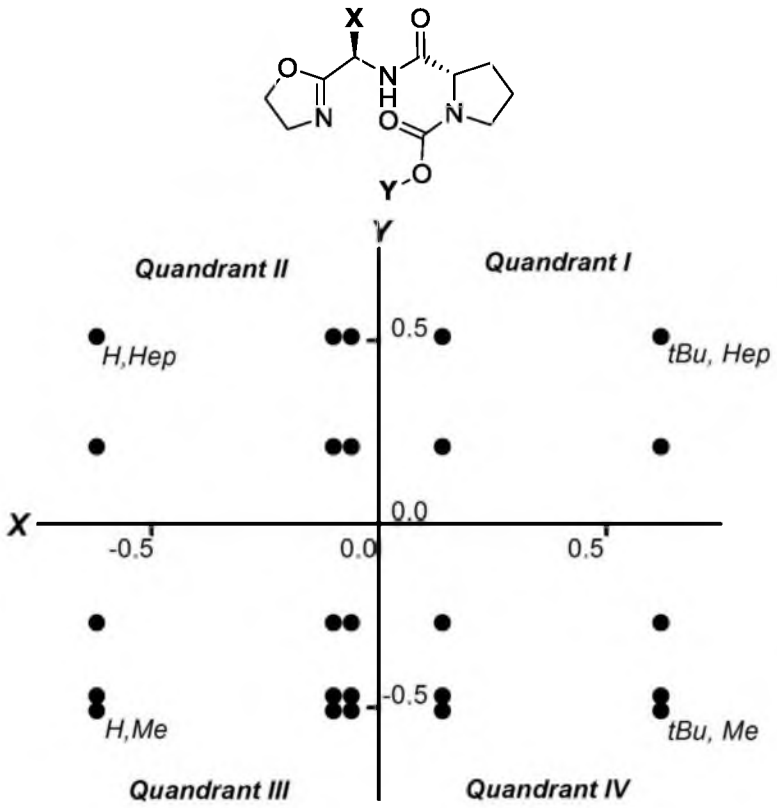


Figure 2.14. Quadrant array of data points for the 5×5 ligand library showing a bias of data in quadrant III.

and Y axes, appeared to be an ideal reconciliation of minimum synthetic effort with potential prediction of better performing catalysts.

The second principle states that extrapolation outside the range of a data spread is far less accurate than interpolation. In order to accommodate this second principle of experimental design, the types of ligands used in the library had to be rethought. To expand our model to make it interpolative instead of extrapolative, we designed a new nearly evenly spread library of 9 ligands, which also employed the previously synthesized ligand **3a**-**c** (Figure 2.15). The bulky CEt_3 substituent has one of the largest Charton values reported. Although synthesis of larger substituents at the Y position is theoretically possible, such substituents could not be quantified because they lack Charton values. Thus, the CEt_3 substituent constitutes the upper bound along the Y axis and the Me substituent the lower bound. On the X axis the largest readily available substituent was the tBu group, which represents the upper bound and hydrogen the lower bound. One result of using the Y_{CEt3} ligands is that the data had to be readjusted using Equation 2.5. Table 2.2 lists the adjusted Charton values for the 3×3 ligand library. The center points chosen were Me and *t*Bu in the X and Y axes respectively. The spread of data points in the X and Y axes is given in Figure 2.15. These center points do not lie perfectly on the center of the spread but, because of the discreet nature of the substituents, are reasonable approximations. The new design encapsulates the available synthetic space with a minimal number of ligands. Of note, the X_{tBu} Y_{CEt3} ligand proved synthetically difficult to prepare and was substituted in the data array for the X_{iPr} Y_{CEt3} ligand. The effect of this is the indicated region where the predictive power of the model becomes poor.

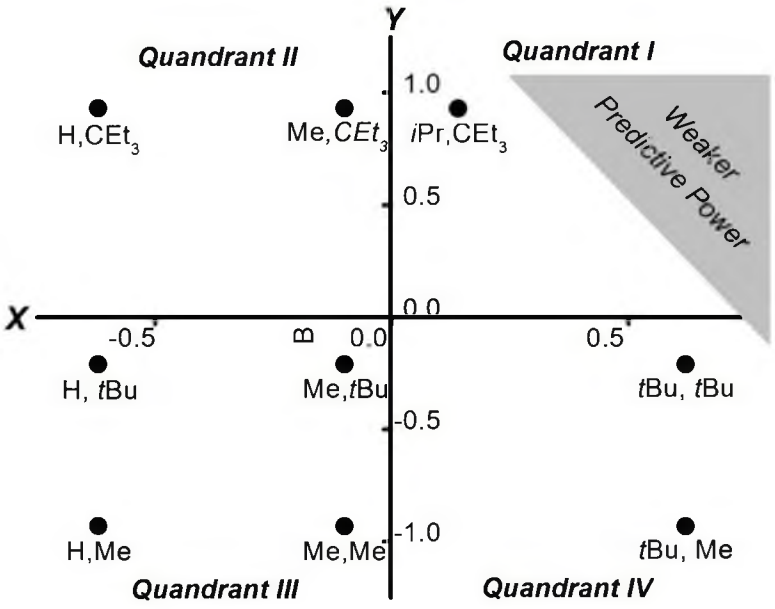


Figure 2.15. Quadrant array of the revised 3×3 ligand library.

Table 2.2. Adjusted Charton values for the 3×3 ligand library.

		Adj. Charton Values	
Substituent		**X**	**Y**
X_{H}	Y_{Me}	-0.62	-0.93
X_{H}	Y_{tBu}	-0.62	-0.21
X_{H}	Y_{CEt3}	-0.62	0.93
X_{Me}	Y_{Me}	-0.1	-0.93
X_{Me}	Y_{tBu}	-0.1	-0.21
X_{Me}	Y_{CEt3}	-0.1	0.93
X_{tBu}	Y_{Me}	0.62	-0.93
X_{tBu}	Y_{tBu}	0.62	-0.21
X_{tBu}	Y_{CEt3}	0.62	0.93

Reevaluation of the Data

Reevaluation of the data for allylation of benzaldehyde also incorporated a new technique for regression. Using the above method to regress the data for benzaldehyde revealed a model described by Equation 2.11 and depicted in Figure 2.16A.

$$\Delta\Delta G^{\ddagger} = 0.92 - 0.53X - 0.89X^2 - 0.89Y^2 - 0.69XY - 0.97XY^2 \quad (2.11)$$

Because the nine-member ligand library was employed to generate the model, the remaining data of the original 25-membered library could be used to externally validate the model. The validation plot of predicted and measured enantioselectivities is given in Fig 2.16B. Q^2 is a common statistical measure for predictive validation, where Q^2 is the slope of a predicted versus measured plot. The model gives a Q^2 value of 0.6 where values > 0.5 are generally considered predictive in the QSAR literature.[22]

The surface model possesses a clear maximum point within the synthetic space. Again because of the discreet nature of Charton values, the model can only be treated in a semi-empirical fashion. The model maximum falls between the values for ligand **4a** and **4b** shown in Figure 2.17. These two ligands were the best performing ligands evaluated with the truncated scaffold in this allylation reaction. In our 3×3 ligand library, both of these ligands were not included in the training set. The model developed from only 9 ligands was able to accurately predict a priori the optimal ligand structure. Prediction of the performance of these ligands is lower than experimentally observed, which might be attributed to the small and unavoidable biases in the data set.

Comparison of Equations 2.10 and 2.11 show the presence of the same crossterms in both models with roughly the same direction and magnitude. The preservation of the crossterms across models derived using different data further validates the proposed synergistic relationship of the X and Y substituents.

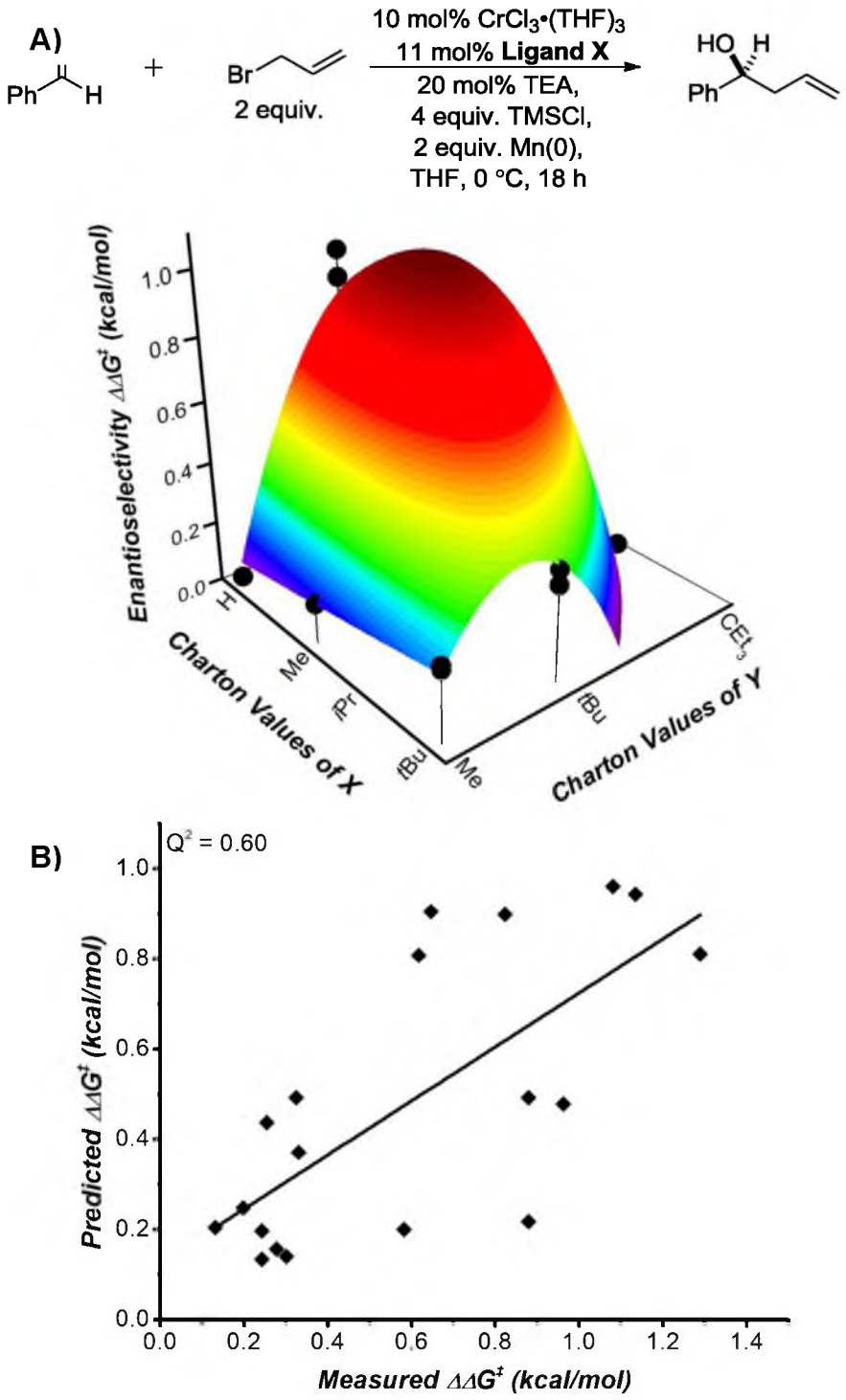
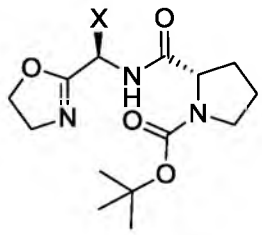


Figure 2.16. A) Surface model given by Equation 11 for the allylation of benzaldehyde under standard conditions. B) Validation plot for 19 other ligands not included in the training data.



	X	*Predicted er*	*Error*	*Measured er*
4a	Et	16:84	±5	9.5:90.5
4b	*i*Pr	17:83	±5	8:92

Figure 2.17. Predictions of the optimal ligand from Equation 11.

The allylation of acetophenone was reevaluated using the revised 3×3 ligand library and data analysis techniques. The revised model is given in Equation 2.12 and the surface is given in Figure 2.18A.

$$\Delta\Delta G^{\ddagger} = 1.31 + 0.046X - 0.69X^2 - 1.1Y^2 - 0.069XY + 0.92Y^3 \quad (2.12)$$

Again, independent validation was performed using ligands not contained in the training set. The Q^2 value shown in Figure 2.18B again indicates the model is reasonably predictive.

The surface described by Equation 2.12 predicts a maximum enantioselectivity achievable for this ligand scaffold. Similar to the above example, this maximum was determined to lie closest to ligand **5** (Figure 2.19). Evaluating all of the data gathered for this reaction and ligand scaffold holds that this ligand is the highest performing of any ligand in the enantioselective ketone allylation reaction. The data for this ligand was not incorporated into the training set again, demonstrating the ability of 3D QSAR to predict optimal structure a priori. The predicted enantioselectivity for **5** was lower than experimentally observed as was the case in the 3×3 allylation of benzaldehyde model.

Both of these examples demonstrate the ability of 3D free energy surfaces in interpolating the optimal ligand structure. However, the models developed were based on systems, which we had studied and reported optimized reaction conditions. Our desire was to use this system on challenging reactions to quickly identify optimal ligand structure for high enantioselectivity. To explore the potential of the system on a challenging reaction, we explored the NHK allylation of methyl ethyl ketone (MEK). Our previous reports of enantioselective NHK reactions had focused on aryl substrates primarily because observed enantioselectivity was poor for aliphatic substrates. We hoped that the development of a 3D free energy surface for the allylation of the simplest aliphatic substrate, MEK, would direct us to an optimized ligand structure for the allylation of aliphatic ketones.

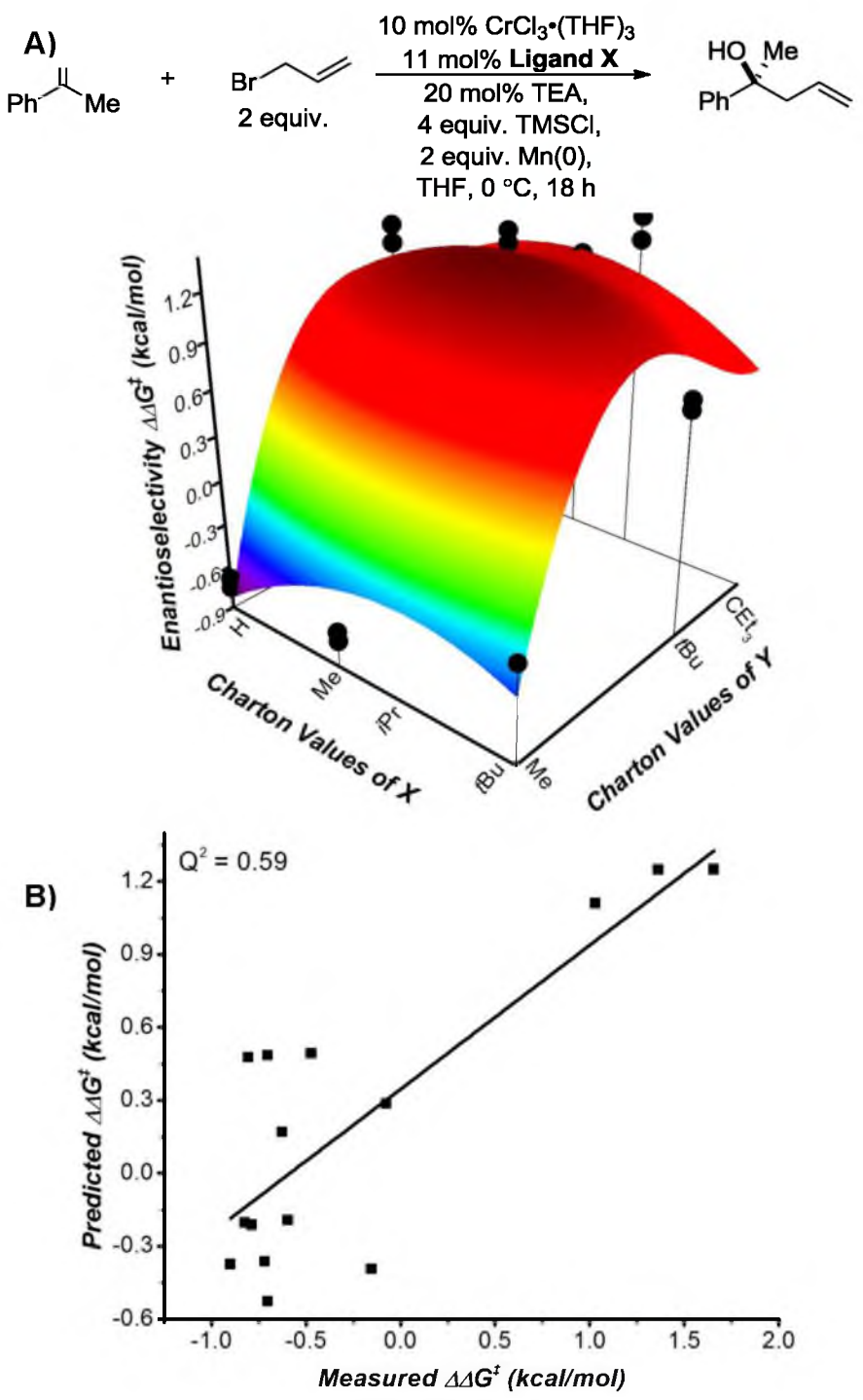
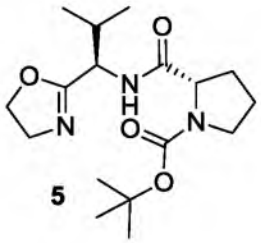


Figure 2.18. A) Surface model given by Equation 12 for that allylation of acetophenone under standard conditions. B) Validation plot for 19 other ligands not included in the training data.



5

X	Y	Predicted er	Error	Measured er
*i*Pr	*t*Bu	8:92	±2	4.5:95.5

Figure 2.19. Predictions of the optimal ligand allylation of acetophenone.

To probe the allylation of MEK, we utilized the same reaction conditions that had been previously used in our allylation reactions. Evaluation of the nine-member ligand library and analysis of the subsequent data yielded Equation 2.13 which is shown in Figure 2.20.

$$\Delta\Delta G^{\ddagger} = -0.87 - 0.94X - 0.29Y + 0.11X^2 + 0.20Y^2 + 0.11XY + 2.17X^3 + 0.44XY^2 \quad (2.13)$$

Disappointingly, the model revealed a rather featureless relationship between enantioselectivity and positions X and Y. Interestingly, the optimal ligand observed in the screen was $X_{Me,}$ Y_{Me} indicating that smaller catalyst features were desirable. However, the conclusion that must be drawn from the model is that modification of the ligand scaffold will not lead to the desired levels of enantioselectivity. The surface indicates major modification to the reaction must be explored to achieve higher enantioselectivity either through condition modification or ligand restructuring. In this case, the entire 25-membered library was not explored in the reaction before this conclusion was drawn. Hence, full validation of the model was not pursued.

Although disappointing, the analysis of the allylation of MEK presents another desirable aspect of this type of analysis. After evaluating only nine ligands, we were able to determine that satisfactory enantioselectivity was not likely to be achieved using the current ligand and conditions. The result was that we could change course if our goal is to obtain a highly enantioselective reaction. Not including synthesis of the ligands, evaluation of the reaction took less than a week. Frequently in asymmetric reaction development, a greater investment of time and ligand synthesis is required before similar conclusions can be drawn.

Conclusion

Through the use of 3D free energy surfaces, we were able to make accurate predictions about catalyst performance and design. After some initial crude attempts led to some predictive power, we were able to apply the principles of experimental design to develop a nine-

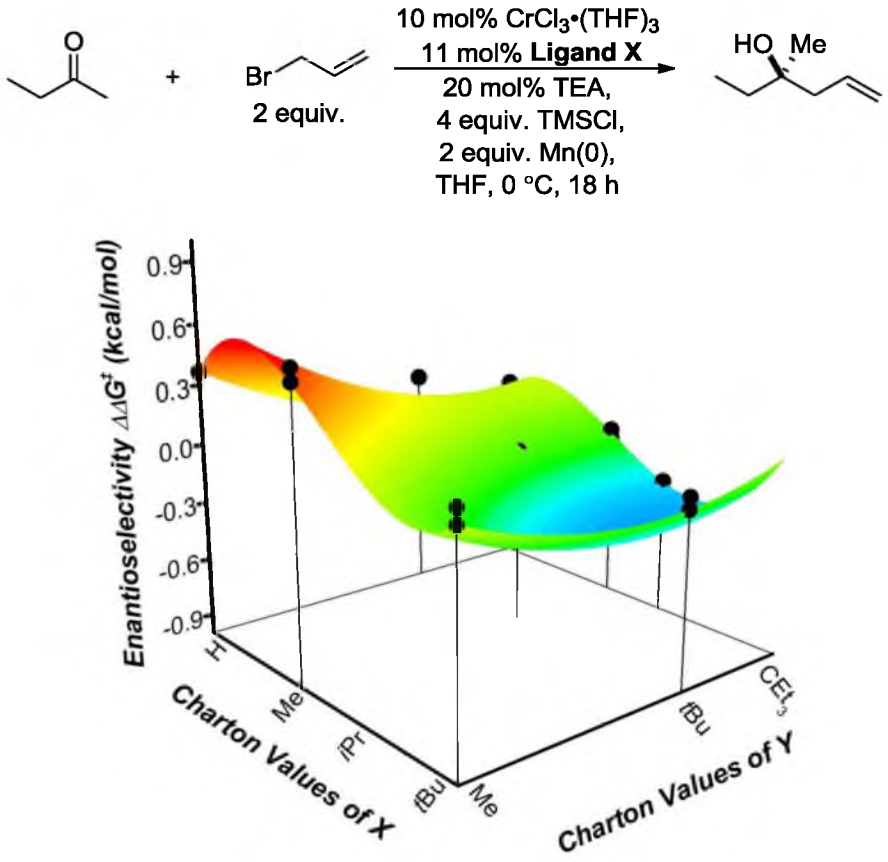


Figure 2.20. Surface model given by Equation 13 for that allylation of MEK under standard conditions.

membered library, which was used to develop models able to correctly predict the optimal ligand structure for the allylation of benzaldehyde as well as acetophenone. Only a handful of examples exist where accuracy of this level was achieved in predicting enantioselective outcomes in asymmetric catalysis.[23-26] Through the use of Charton's steric parameters, evidence was found via statistically significant crossterms that both the X and Y positions were interacting synergistically. This hypothesis would have been difficult to examine via other techniques. The overall approach does not require kinetic or structural data beyond the substituents examined.

This study demonstrates the power of experimental design applied to catalysis or reaction optimization. Although this study focused on steric-steric-enantioselective relationships, the techniques are not limited to such relationships or dimensionality. Any quantifiable, variable catalyst characteristic could be employed as the dependent variables and several measurable quantities could be used in place of enantioselectivity. As we continue to explore the application of these predictive techniques, we hope to expand the scope of independent as well as dependent variables used in them.

Experimental

General Considerations

Unless otherwise noted, all reactions were performed under a nitrogen atmosphere with stirring. Toluene, dichloromethane, dichloroethane and THF were dried before use by passing through a column of activated alumina. Methanol was distilled from magnesium methoxide. Triethylamine was distilled from CaH_2. Benzaldehyde was purified by aqueous base wash, drying with sodium sulfate, and followed by fractional distillation. Acetophenone was purified by drying over Na_2SO_4 then fractional distillation. Methyl Ethyl Ketone was purified by two sequential fractional distillations. Allyl bromide was purified by drying over magnesium

sulfate, filtration and fractional distillation. $CrCl_3(THF)_3$ was prepared by soxhlet extraction of anhydrous $CrCl_3$ with anhydrous THF. All other reagents were purchased from commercial sources and used without further purification. Yields were calculated for material judged homogeneous by thin-layer chromatography and NMR. Thin-layer chromatography was performed with EMD silica gel 60 F254 plates eluting with the solvents indicated, visualized by a 254 nm UV lamp, and stained either with potassium permanganate, phosphomolybdic acid, or ninhydrin. Flash column chromatography was performed with EcoChrom MP Silitech 32-63D 60Å silica gel, slurry packed with solvents indicated in glass columns. Nuclear magnetic resonance spectra were acquired at 300, 400, or 500 MHz for ^{1}H, and 75, 100, or 125 MHz for ^{13}C and 50°C. Chemical shifts for proton nuclear magnetic resonance (^{1}H NMR) spectra are reported in parts per million downfield relative to the line of $CHCl_3$ singlet at 7.24 ppm. Chemical shifts for carbon nuclear magnetic resonance (^{13}C NMR) spectra are reported in parts per million downfield relative to the center-line of the $CDCl_3$ triplet at 77.23 ppm. The abbreviations s, d, t, p, sep, dd, td, bs, and m stand for the resonance multiplicities singlet, doublet, triplet, pentet, septet, doublet of doublets, triplet of doublets, broad singlet, and multiplet, respectively. Optical rotations were obtained (Na D line) using a Perkin Elmer Model 343 Polarimeter fitted with a micro cell with a 1 dm path length. Concentrations are reported in g/100 mL. SFC (super critical fluid chromatography) analysis was performed at 25 °C or 40 °C, using a Thar instrument fitted with chiral stationary phase (as indicated). Melting points were obtained on an electrothermal melting point apparatus and are uncorrected. Unless otherwise noted, glassware for all reactions was oven-dried at 110 °C and cooled in a dry atmosphere prior to use.

Ligand Synthesis and Characterization [5,7,8,27]

Anderson Coupling with Proline (Figure 2.21)

To a flame-dried round bottom flask flushed with N_2 containing a stirbar and fitted with a septum was added the differentially protected proline (1 equiv.) as a solid or as a standard solution in DCM. The flask is fitted with a septum and attached to a positive pressure of N_2. The starting material was then diluted with solvent (0.2M). NMM (1.2 equiv.) was added dropwise and the solution cooled to ~ 0 °C in an ice/water bath. After cooling for 10 min, the IBCF is added slowly dropwise (1 mL/min) and formation of a colorless precipitate is observed. The reaction is allowed to stir 30 min at ~ 0 °C under nitrogen, after which a second portion of NMM (1.1 equiv.) is added. Immediately following the second addition of NMM, the septum is removed and the amino acid methylester hydrochloride salt is added (1.2 equiv.) in one-third portions over 1-2 min. Once the addition of the amino acid methyl ester is complete the ice/water bath is removed and the reaction allowed to stir until starting material is no longer observed by TLC (5% MeOH/DCM, PMA with charring). Upon completion the reaction is diluted in DCM (0.01M) and washed with aqueous sodium bicarbonate. The aqueous layer is then extracted with DCM (3 × 20 mL) and the organic extracts combined and dried with Na_2SO_4 and concentrated and purified by column chromatography.

Thermal Amide Bond Formation (Figure 2.22)

To a flame-dried round bottom flask with stir bar, flushed with N_2 was added to the methyl ester starting material (1 equiv.) as a solid. The starting material was diluted with THF (1M) and toluene (1M). Then 2-aminoethanol (5 equiv.) was added through the septum and fitted with a condenser. The septum was replaced at the top of the condenser and the mixture was stirred in a sand bath in excess of 100 °C for 1-5 days until complete by TLC (50%

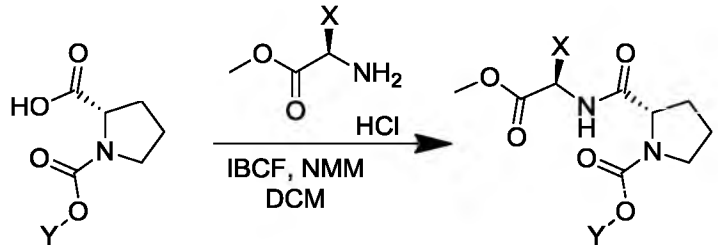


Figure 2.21. Peptide bond formation using Anderson's conditions.

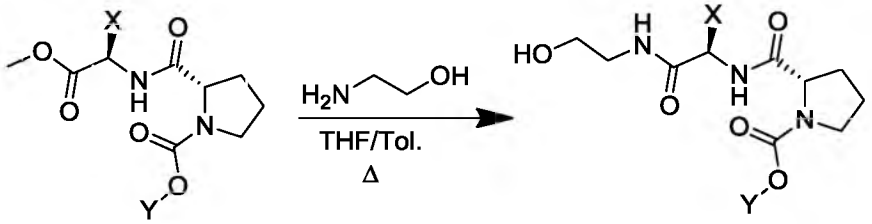


Figure 2.22. Thermal amide-bond formation.

Acetone / Hexanes). Upon completion the reaction was diluted in DCM (0.05M) and washed with an equal volume of water. The aqueous layer was then extracted with DCM (3×20 mL) and the organic extracts combined and dried with Na_2SO_4 and concentrated and purified by column chromatography.

Oxazoline Cyclization (Figure 2.23)

To a flame-dried round bottom flask with stir bar, flushed with N_2 and fitted with a septum was added the amide-alcohol starting material (1 equiv.). The starting material was then diluted in DCM (0.125M) and cooled to ~ -78 °C in an isopropyl/dry ice bath. After 10 min of cooling, the DAST is added dropwise. The reaction is allowed to stir at -78 °C for 1 h and then solid K_2CO_3 is added all at once. The reaction is allowed to warm to ambient temperature and then concentrated. The crude mixture is purified by column chromatography to furnish the ligand products.

(S)-methyl-2-((4,5-dihydrooxazol-2-yl)methylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization, was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.301 g. R_f = 0.3 w/ 10% MeOH/DCM, yellow oil, $[\alpha]^{20}_D$ = –69.5° (c = 0.965, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.99 (bs, 1 H), 4.16 (s, 1 H), 4.12 (s, 1H), 3.855 (d, *J* = 4, 2 H), 3.65 (td, *J* = 9.6, 1.6; 2 H), 3.53 (s, 3 H), 3.34 (bs, 2 H), 2.1-1.65 (m, 4 H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 171.9, 164.5, 67.9, 60.5, 54.0, 52.5, 46.9, 36.7, 28.9, 24.0. HRMS $C_{11}H_{17}N_3O_4$ $(M+H)^+$ calcd. 256.1297, obsvd. 256.1300.

(S)-ethyl-2-((4,5-dihydrooxazol-2-yl)methylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.150 g of **L2.** R_f = 0.4 w/ 10% MeOH/DCM, yellow oil, $[\alpha]^{20}_D$ = –69.7° (c = 1.095, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 7.02 (bs, 1H), 4.16 (s, 1H), 4.11 (t, *J* = 9.6,

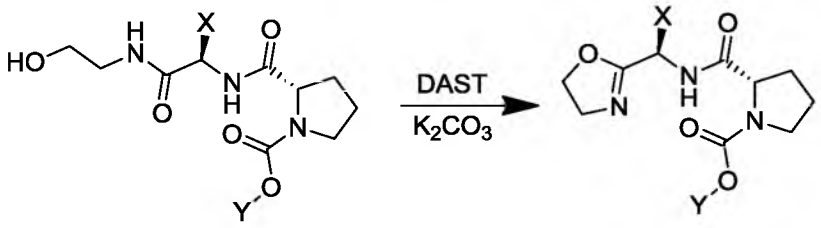


Figure 2.23. Oxazoline cyclization.

2H), 3.96 (m, 2H), 3.85 (s, 2H), 3.65 (dt, *J* = 9.6,1.6; 2H), 3.33 (bs, 2H), 2.1-1.65 (m, 4H), 1.071 (t, *J* = 7.2, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 172.0, 164.5, 67.9, 61.3, 60.4, 54.0, 46.9, 36.7, 28.9, 24.0, 14.3. HRMS $C_{12}H_{19}N_3O_4$ $(M+H)^+$ calcd. 270.1454, obsvd. 270.1451.

(S)-isopropyl-2-((4,5-dihydrooxazol-2-yl)methylcarbamoyl)pyrrolidine-1-carboxylate.

The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.190 g. R_f = 0.4 w/ 10% MeOH/DCM, yellow oil, $[\alpha]^{20}_D$ = – 0.4° (c = 0.055, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 7.08 (bs, 1H), 4.81 (sep, *J* = 6.4, 1H), 4.17 (m, 3H), 3.92 (d, *J* = 4.8, 2H), 3.71 (dt, *J* = 9.6, 1.6; 2H), 3.37 (bs, 2H), 2.2-1.7 (m, 4H), 1.13 (m, 6H), ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 172.1, 164.6, 69.0, 68.1, 60.5, 54.1, 47.0, 36.9, 28.5, 24.1, 22.1. HRMS $C_{13}H_{21}N_3O_4$ $(M+H)^+$ calcd. 284.1610, obsvd. 284.1606.

(S)-tert-butyl-2-((4,5-dihydrooxazol-2-yl)methylcarbamoyl)pyrrolidine-1-carboxylate.

The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.167 g. R_f = 0.4 w/ 10% MeOH/DCM, yellow oil, $[\alpha]^{20}_D$ = – 11.8° (c = 0.195, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.63 (bs, 1H), 4.23 (m, 1H), 4.23 (t, *J* = 9.2, 2H), 3.99 (m, 2H), 3.77 (t, *J* = 9.6, 2H), 3.40 (bs, 2H), 2.29-1.75 (bs, 4H), 1.41 (s, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 172.5, 164.8, 80.6, 68.3, 60.7, 54.3, 47.2, 37.1, 28.5, 24.3. HRMS $C_{14}H_{23}N_3O_4$ $(M+H)^+$ calcd. 298.1767, obsvd. 298.1765.

(S)-heptan-4-yl 2-((4,5-dihydrooxazol-2-yl)methylcarbamoyl)pyrrolidine-1-carboxylate.

The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.175 g. R_f = 0.6 w/ 10% MeOH/DCM, yellow oil, $[\alpha]^{20}_D$ = – 9.7° (c = 0.230, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.91 (bs, 1H), 4.75 (m, 1H), 4.23 (m, 3H), 3.98 (s, 2H), 3.77 (t, *J* = 9.6, 2H), 3.45 (bs, 2H), 2.3-1.78 (m, 4H), 1.47 (m, 4H), 1.29 (m 4H), 0.85 (m, 6H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 167.3, 75.7, 68.3, 60.7, 54.3, 47.2, 37.1, 36.7, 24.4, 18.7, 18.6, 14.1. HRMS $C_{17}H_{29}N_3O_4$ $(M+H)^+$ calcd. 340.2236, obsvd. 340.2239.

(S)-methyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)ethylcarbamoyl)pyrrolidine-1-carboxylate.

The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.365 g. R_f = 0.4 w/ 10% MeOH/DCM, colorless solid, MP = 102-105 °C . $[\alpha]^{20}_D$ = –9.2° (c = 0.140, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.91 (bs, 1H), 4.61 (p, *J* = 6.8, 1H), 4.23 (m, 3H), 3.77 (t, *J* = 10, 2H), 3.66 (s, 3H), 3.47 (bs, 2H), 2.2-1.75 (m, 4H), 1.33 (d, *J* = 7.2, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 171.6, 168.4, 68.4, 61.1, 54.3, 52.8, 47.4, 43.8, 30.2, 24.2, 19.1. HRMS $C_{12}H_{19}N_3O_4$ $(M+H)^+$ calcd. 270.1454, obsvd. 270.1453.

(S)-ethyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)ethylcarbamoyl)pyrrolidine-1-carboxylate.

The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.278 g**.** R_f = 0.4 w/ 10% MeOH/DCM, colorless solid, MP = 93-94 °C . $[\alpha]^{20}_D$ = –13.8° (c = 0.225, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.86 (bs, 1H), 4.57 (p, *J* = 11.2, 1H), 4.19 (t, *J* = 9.6, 2H), 4.06 (m, 1H), 3.73 (t, *J* = 9.6, 2H), 3.43 (bs, 2H), 2.20-1.70 (m, 4H), 1.29 (d, *J* = 7.2, 3H), 1.17 (m, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ =171.5, 168.3, 68.3, 61.6, 60.9, 54.2, 47.2, 43.6, 29.7, 24.1, 19.1, 14.7. HRMS $C_{13}H_{21}N_3O_4$ $(M+H)^+$ calcd. 284.1610, obsvd. 284.1609.

(S)-isopropyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)ethylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.285 g. R_f = 0.5 w/ 10% MeOH/DCM, colorless solid, MP = 79-80 °C . $[\alpha]^{20}_D$ = –9.5° (c = 0.165, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.65 (bs, 1H), 4.88 (sep, *J* = 6, 1H), 4.62 (p, *J* = 7, 1H), 4.22 (m, 3H), 3.76 (t, *J* = 9.6, 2H), 3.45 (bs, 2H), 2.30-1.75 (m, 4H), 1.33 (d, *J* = 6.8, 3H), 1.19 (t, *J* = 6.8, 6H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 171.7, 168.4, 69.2, 68.4, 61.0, 54.3, 47.2, 43.7, 29.8, 24.1, 22.3, 22.3, 19.2. HRMS $C_{14}H_{23}N_3O_4$ $(M+H)^+$ calcd. 298.1767, obsvd. 298.1765.

(S)-tert-butyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)ethylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.325 g. R_f = 0.5 w/ 10% MeOH/DCM, colorless solid, MP = 91-94 °C . $[\alpha]^{20}_D$ = –14.4° (c = 0.195, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.65 (bs, 1H), 4.63 (p, *J* = 8, 1H), 4.22 (m, 3H), 3.76 (t, *J* = 8.8, 2H), 3.42 (bs, 2H), 2.3-1.73 (m, 4H), 1.41 (s, 9H), 1.35 (dd, *J* = 4.8, 2; 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 174.1, 168.3, 80.5, 68.4, 54.3, 47.2, 43.7, 28.5, 24.1, 19.3. HRMS $C_{15}H_{25}N_3O_4$ $(M+H)^+$ calcd. 312.1923, obsvd. 312.1922.

(S)-heptan-4-yl-2-((R)-1-(4,5-dihydrooxazol-2-yl)ethylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.125 g. R_f = 0.5 w/ 10% MeOH/DCM, colorless solid, MP = 91-93 °C . $[\alpha]^{20}_D$ = –5° (c = 0.090, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.68 (bs, 1H), 4.80 (m, 1H), 4.65 (p, *J* = 7.2, 1H), 4.25 (m, 3H), 3.79 (t, *J* = 10.4, 2H), 3.49 (bs, 2H), 2.35-1.78 (m, 4H), 1.60-1.25 (m, 8H), 0.89 (t, *J* = 7.2, 6H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 168.5, 164.3, 75.8, 68.5, 61.1, 54.4, 47.3, 43.8, 36.8, 19.2, 18.8, 14.2, 14.2. HRMS $C_{17}H_{29}N_3O_4$ $(M+H)^+$ calcd. 354.2393, obsvd. 354.2390.

(S)-methyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)propylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.213 g. R_f = 0.5 w/ 10% MeOH/DCM, colorless solid, MP = 109-110 °C . $[\alpha]^{20}_D$ = –8.2° (c = 0.125, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.65 (bs, 1H), 4.60 (q, *J* = 6.8, 1H), 4.25 (m, 3H), 3.79 (t, *J* = 9.6, 2H), 3.69 (s, 3H), 3.50 (bs, 2H), 2.20-1.80 (m, 4H), 1.66 (sep, *J* = 7.2, 1H), 0.86 (t, *J* = 7.6, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 167.5, 68.3, 61.2, 54.4, 52.9, 48.8, 47.5, 29.8, 26.1, 24.9, 24.2, 9.3. HRMS $C_{13}H_{21}N_3O_4$ $(M+H)^+$ calcd. 284.1610, obsvd. 284.1609.

(S)-ethyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)propylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.315 g. R_f = 0.5 w/ 10% MeOH/DCM, colorless solid, MP = 99-101 °C . $[\alpha]^{20}_D$ = –2.9° (c = 0.050, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.65 (bs, 1H), 4.61 (q, *J* = 6.8, 1H), 4.24 (m, 3H), 4.13 (q, *J* = 6.8, 1H), 3.79 (t, *J* = 9.6, 2H), 3.50 (bs, 2H), 2.25-1.79 (m, 4H), 1.66 (sep, *J* = 7.2, 1H), 1.23 (t, *J* = 7.2, 3H), 0.869 (t, *J* = 7.6, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 165.4, 68.3, 61.8, 61.1, 54.4, 48.8, 47.4, 26.2, 24.3, 14.8, 9.4. HRMS $C_{14}H_{23}N_3O_4$ $(M+H)^+$ calcd. 298.1767, obsvd. 298.1769.

(S)-isopropyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)propylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.190 g. R_f = 0.5 w/ 10% MeOH/DCM, colorless solid, MP = 92-93 °C . $[\alpha]^{20}_D$ = –5.1° (c = 0.09, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.67 (bs, 1H), 4.92 (sep, *J* = 10.4, 1H), 4.69 (q, *J* = 6.8, 1H), 4.31-4.21 (m, 3H), 3.80 (t, *J* = 9.6, 2H), 3.49 (bs, 2H), 2.30-1.80 (m, 4H), 1.652 (sep, *J* = 8, 1H), 1.22 (t, *J* = 5.6, 3H), 0.89 (t, *J* = 7.2, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 170.9, 167.5, 69.3, 68.3, 61.1, 54.4, 48.9, 47.4, 26.3, 24.2, 22.4, 22.3, 9.5. HRMS $C_{15}H_{25}N_3O_4$ $(M+H)^+$ calcd. 312.1923, obsvd. 312.1921.

(S)-tert-butyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)propylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 5-10% MeOH/DCM as eluent to give 0.330 g. R_f = 0.5 w/ 10% MeOH/DCM, colorless solid, MP = 131-132 °C . $[\alpha]^{20}_D$ = –4.5° (c = 0.080, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.98 (bs, 1H), 4.61 (q, *J* = 7.2, 1H), 4.24 (m, 3H), 3.79 (t, *J* = 9.6, 2H), 3.45 (bs, 2H), 2.30-1.80 (m, 4H), 1.68 (sep, *J* = 7.2, 1H), 1.44 (s, 9H), 0.89 (t, *J* = 7.2, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 172.7, 167.5, 94.7, 68.2, 61.1, 54.4, 48.8, 47.3, 28.6, 26.3, 24.1, 9.5. HRMS $C_{16}H_{27}N_3O_4$ $(M+H)^+$ calcd. 326.2080, obsvd. 326.2079.

(S)-heptan-4-yl-2-((R)-1-(4,5-dihydrooxazol-2-yl)propylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 1-10% MeOH/DCM as eluent to give 0.198 g. R_f = 0.6 w/ 50% Acetone/Hexanes, colorless solid, MP = 74-75 °C . $[\alpha]^{20}_D$ = –6.1° (c = 0.130, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.67 (bs, 1H), 4.79 (m, 1H), 4.59 (q, *J* = 7.2, 1H), 4.31-4.20 (m, 3H), 3.78 (t, *J* = 9.6, 2H), 3.49 (bs, 2H), 2.32-1.78 (m, 4H), 1.65 (sep, *J* = 7.2, 1H), 1.49 (m, 4H), 1.30 (m, 4H), 0.87 (m, 6H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 174.6, 167.5, 75.7, 68.3, 61.0, 54.6, 48.9, 47.4, 36.8, 36.8, 26.3, 24.1, 18.8, 18.7, 14.2, 14.1, 9.5. HRMS $C_{19}H_{33}N_3O_4$ $(M+H)^+$ calcd. 368.2549, obsvd. 368.2552.

(S)-methyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)-2-methylpropylcarbamoyl)pyrrolidine-1-carboxylate. Known compound.[5]

(S)-ethyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)-2-methylpropylcarbamoyl)pyrrolidine-1-carboxylate. Known compound.[5]

(S)-isopropyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)-2-methylpropylcarbamoyl)pyrrolidine-1-carboxylate. Known compound.[7]

(S)-tert-butyl-2-((R)-1-(4,5-dihydrooxazol-2-yl)-2-methylpropylcarbamoyl)pyrrolidine-1-carboxylate. Known compound[5]

(S)-heptan-4-yl-2-((R)-1-(4,5-dihydrooxazol-2-yl)-2-methylpropylcarbamoyl)pyrolidine-1-carboxylate. The product of the DAST cyclization was purified flash silica-gel column chromatography with 1-10% MeOH/DCM as eluent to give 0.300 g. R_f = 0.7 w/ 10% MeOH/DCM, colorless solid, MP = 75-77 °C . $[\alpha]^{20}_D$ = –3.6° (c = 0.080, $CHCl_3$). ^{1}H-NMR (300 MHz, $CDCl_3$) δ = 6.86 (bs, 1H), 4.80 (td, *J* = 7.2, 5.5; 1H), 4.57 (q, *J* = 5.1, 1H), 4.34 (d, *J* = 5.6, 1H), 4.22 (td, *J* = 9, 3; 2H), 3.79 (t, *J* = 9.6, 2H), 3.50 (m, 2H), 2.40-1.80 (m, 4H), 1.50 (m, 4H), 1.33 (m, 4H), 0.90 (m, 6H). ^{13}C-NMR {^{1}H} (75 MHz, $CDCl_3$) δ = 172.8, 167.1, 75.8, 68.1, 61.1, 54.3, 52.7, 47.3,

36.9, 36.8, 31.8, 18.9, 18.8, 18.7, 17.9, 14.2, 14.2. HRMS $C_{20}H_{35}N_3O_4$ $(M+H)^+$ calcd. 382.2706, obsvd. 382.2708.

(S)-methyl-2-((S)-1-(4,5-dihydrooxazol-2-yl)-2,2-dimethylpropylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 50% Acetone/Hexane as eluent to give 0.140 g. R_f = 0.4 w/ 50% Acetone/Hexanes, colorless solid, MP = 133-135 °C. $[\alpha]^{20}_D$ = +3.4° (c = 0.055, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.71 (bs, 1H), 4.52 (d, *J* = 8, 1H), 4.21 (m, 3H), 3.79 (m, 2H), 3.69 (s, 3H), 3.51 (bs, 2H), 2.20-1.80 (m, 4H), 0.94 (s, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 171.5, 166.8, 67.7, 61.3, 55.5, 54.3, 52.9, 47.5, 35.5, 26.7, 24.2. HRMS $C_{15}H_{25}N_3O_4$ $(M+H)^+$ calcd. 312.1923, obsvd. 312.1924.

(S)-ethyl-2-((S)-1-(4,5-dihydrooxazol-2-yl)-2,2-dimethylpropylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 50% Acetone/Hexane as eluent to give 0.320 g. R_f = 0.45 w/ 10% Acetone/Hexanes, colorless solid, MP = 134-136 °C . $[\alpha]^{20}_D$ = +5.8° (c = 0.105, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.73 (bs, 1H), 4.53 (d, *J* = 9.6, 1H), 4.33 (d, *J* = 7.2, 1H), 4.17 (m, 4H), 3.80 (m, 2H), 3.51 (bs, 2H), 2.25-1.80 (m, 4H), 1.24 (t, *J* = 7.2, 3H), 0.96 (s, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 180.0, 166.8, 67.7, 61.9, 61.2, 55.2, 54.4, 47.43, 41.4, 26.7, 24.2, 14.9. HRMS $C_{16}H_{27}N_3O_4$ $(M+H)^+$ calcd. 326.2080, obsvd. 326.2078.

(S)-isopropyl-2-((S)-1-(4,5-dihydrooxazol-2-yl)-2,2-dimethylpropylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 50% Acetone/Hexane as eluent to give 0.290 g. R_f = 0.55 w/ 50% Acetone/Hexanes, colorless solid, MP = 148-150 °C . $[\alpha]^{20}_D$ = +5.4° (c = 0.090, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.98 (bs, 1H), 4.92 (sep, *J* = 6.4, 1H), 4.53 (d, *J* = 9.6, 1H), 4.58 (d, *J* = 6.8, 1H), 4.23 (m, 2H), 3.79 (m, 2H), 3.49 (bs, 2H), 2.28-1.80 (m, 4H), 1.23 (d, *J* =

6, 6H), 0.96 (s, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 171.3, 166.8, 94.7, 69.4, 67.6, 61.1, 55.5, 54.5, 47.4, 35.5, 26.8, 22.4, 22.4. HRMS $C_{17}H_{29}N_3O_4$ $(M+H)^+$ calcd. 340.2236, obsvd. 340.2242.

(S)-tert-butyl-2-((S)-1-(4,5-dihydrooxazol-2-yl)-2,2-dimethylpropylcarbamoyl)pyrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 50% Acetone/Hexane as eluent to give 0.300 g. R_f = 0.5 w/ 10% Acetone/Hexanes, colorless solid, MP = 141-143 °C . $[\alpha]^{20}_D$ = +6.1° (c = 0.095, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.97 (bs, 1H), 4.52 (d, *J* = 9.6, 1H), 4.21 (m, 3H), 3.82 (m, 2H), 3.44 (bs, 2H), 2.25-1.80 (m, 4H), 1.44 (s, 9H), 0.96 (s, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 166.8, 67.6, 61.4, 55.5, 54.3, 47.3, 41.9, 35.4, 28.6, 26.7, 24.1. HRMS $C_{18}H_{31}N_3O_4$ $(M+H)^+$ calcd. 354.2392, obsvd. 354.2390.

(S)-heptan-4-yl-2-((S)-1-(4,5-dihydrooxazol-2-yl)-2,2-dimethylpropylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 50% Acetone/Hexane as eluent to give 0.210 g. R_f = 0.6 w/ 10% MeOH/DCM, colorless solid, MP = 103-105 °C. $[\alpha]^{20}_D$ = +2.7° (c = 0.070, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.99 (bs, 1H), 4.75 (m, 1H), 4.47 (t, *J* = 9.6), 1H), 4.29 (m, 1H), 4.15 (m, 2H), 3.75 (m, 2H), 3.44 (bs, 2H), 2.3-1.75 (m, 4H), 1.46 (m, 4H), 1.28 (m, 4H), 0.89 (m, 9H), 0.83 (m, 6H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 172.1, 166.7, 75.7, 67.5, 66.9, 55.5, 54.2, 47.2, 36.8, 36.7, 35.3, 26.7, 24.1, 18.7, 18.7, 14.1. HRMS $C_{21}H_{37}N_3O_4$ $(M+H)^+$ calcd. 396.2862, obsvd. 396.2859.

(S)-1-(3,3-diethylpentanoyl)-N-((4,5-dihydrooxazol-2-yl)methyl)pyrrolidine-2-carboxamide. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 1-5% MeOH/DCM as eluent to give 0.182 g. R_f = 0.6 w/ 10% MeOH/DCM, yellow oil. $[\alpha]^{20}_D$ = -34.7° (c = 0.11, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.67 (bs, 1H), 4.627 (m, 1H), 4.23 (m, 3H), 3.77 (t, *J* = 9.2, 2H), 3.43 (bs, 2H), 2.23-1.8 (m, 4H), 1.79 (q, *J* = 7, 4H), 1.34

(d, *J* = 7.2, 3H), 0.783 (t, *J* = 7.2, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 168.4, 68.4, 60.5, 54.4, 47.3, 47.2, 43.7, 28.3, 27.4, 24.2, 19.2, 7.8. HRMS $C_{21}H_{37}N_3O_4$ $(M+H)^+$ calcd. 340.2236, obsvd. 340.2237.

(S)-3-ethylpentan-3-yl-2-((R)-1-(4,5-dihydrooxazol-2-yl)ethylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 1-5% MeOH/DCM as eluent to give 0.304 g. R_f = 0.7 w/ 10% MeOH/DCM, colorless solid, MP = 88-90 °C. $[\alpha]^{20}_D$ = -44.0° (c = 0.310, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.68 (bs, 1H), 4.63 (m, 1H), 4.23 (m, 3H), 3.77 (t, *J* = 9.2, 2H), 3.43 (bs, 2H), 2.23-1.80 (m, 4H), 1.79 (q, *J* = 7, 6H), 1.34 (d, *J* = 7.2, 3H), 0.78 (t, *J* = 7, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 168.4, 68.4, 60.5, 54.4, 47.3, 43.7, 28.3, 27.4, 24.2, 19.2, 7.8. HRMS $C_{21}H_{37}N_3O_4$ $(M+H)^+$ calcd. 354.2393, obsvd. 354.2396.

(S)-3-ethylpentan-3-yl-2-((R)-1-(4,5-dihydrooxazol-2-yl)propylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 1-5% MeOH/DCM as eluent to give 0.252 g. R_f = 0.7 w/ 10% MeOH/DCM, colorless solid, MP = 99-103 °C. $[\alpha]^{20}_D$ = -35.9° (c = .350, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.74 (bs, 1H), 4.58 (m, 1H), 4.22 (m, 3H), 3.77 (t, *J* = 9.6, 2H), 3.44 (bs, 2H), 2.32-1.80 (m, 4H), 1.80 (q, *J* = 7.2, 6H), 0.87 (m, 3H), 0.79 (t, *J* = 7.2, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 167.5, 68.2, 54.4, 48.8, 43.7, 47.2, 27.4, 26.3, 24.4, 9.6, 7.9. HRMS $C_{21}H_{37}N_3O_4$ $(M+H)^+$ calcd. 368.2548, obsvd. 368.2549.

(S)-3-ethylpentan-3-yl-2-((R)-1-(4,5-dihydrooxazol-2-yl)-2-methylpropylcarbamoyl)pyrrolidine-1-carboxylate. The product of the DAST cyclization was purified by flash silica-gel column chromatography with 1-5% MeOH/DCM as eluent to give 0.252 g. R_f = 0.7 w/ 10% MeOH/DCM, colorless solid, MP = 87-88 °C. $[\alpha]^{20}_D$ = 37.0° (c = 0.275, $CHCl_3$). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 6.76 (bs, 1H), 4.56 (m, 1H), 4.30 (d, *J* = 7.6, 1H), 4.22 (m,

2H), 3.77 (t, *J* = 9.2, 2H), 3.44 (bs, 2H), 2.30-1.90 (m, 5H), 1.82 (q, *J* = 7.2, 6H), 0.91 (m, 6H), 0.81 (t, *J* = 7.2, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 170.9, 167.0, 68.0, 54.3, 52.8, 47.4, 47.3, 31.8, 27.5, 19.0, 18.1, 8.0. HRMS $C_{21}H_{37}N_3O_4$ $(M+H)^+$ calcd. 382.2706, obsvd. 382.2710.

Allylation of Benzaldehyde

To an oven dried 1.5 dram vial with stir bar, was added $CrCl_3(THF)_3$ (9.4 mg, 0.025mmol, 0.1 eq.), manganese powder (27.5 mg, 0.5 mmol, 2 eq.), and ligand (0.0275 mmol, 0.11 eq.). The vial was then fitted with a Tefton septum and cap. The vial was further sealed by wrapping the cap with Teflon tape. The vial was then purged under vacuum and flushed with argon gas three times. Under an atmosphere of argon, a standard solution of THF/TEA (0.6% TEA/THF, or 7.5 *u*L per 1.25 ml, in the test reaction 1.25 ml of this solution was added to each vial) was added. The reaction was then stirred for 3-5 min and TMSCl (124 *u*L, 1 mmol, 4 eq.) was added dropwise. The reaction was then allowed to stir for 20 min at RT. After 20 min, the reaction had turned dark grey. The vial was transported to a cold bath at 0 °C and allowed to equilibrate for 10 min After equilibration, allyl bromide (45 *u*L, 0.5 mmol, 2 eq.) was added. The reaction was then allowed to stir for another 20 min and finally benzaldehyde was added. The reaction was then stirred vigorously for 20 h and slowly turned light gray or greenish gray. The reaction was quenched by slow addition of a saturated sodium bicarbonate solution (approx. 1 mL). The mixture was separated and the aqueous layer washed with diethyl ether (1x1ml). The organic layers were combined and passed over a Celite plug and collected. The organic layer was then dried with sodium sulfate and passed over a silica plug. The silica was rinsed with diethyl ether (~1mL). Samples of the corresponding enantiomeric products were then prepped for analysis without consideration for yield.

(R)-1-phenylbut-3-en-1-ol. Known Compound.[7] Enantiomeric analysis for (R)-1-phenylbut-3-en-1-ol was carried using a Thar SFC system. Concentrations of approximately 5-10 mg/mL sample were loaded. Injections were made at 5 uL sample loading on a Chiralcel OJ-H (0.46cm x 25 cm) column. The chiral separation was carried out using 1% MeOH at 40 °C. Peak 1 R_T: 7.43 min, Peak 2 R_T: 8.11.

Allylation of Acetophenone

To an oven dried 1.5 dram vial with stir bar, was added $CrCl_3(THF)_3$ (9.4 mg, 0.025mmol, 0.1 equiv.), 325 mesh manganese powder (27.5 mg, 0.5 mmol, 2 eq.), and ligand (.0275 mmol, 0.11 eq.). The vial was then fitted with a Tefton septum and cap. The vial was further sealed by wrapping the cap with Teflon tape. The vial was then purged under vacuum and flushed with argon gas three times. Under an atmosphere of argon, a standard solution of THF/TEA (0.6% TEA/THF, or 7.5 *u*L per 1.25 ml, in the test reaction 1.25 ml of this solution was added to each vial) was added. The reaction was then stirred for 3-5 min and TMSCl (124 *u*L, 1 mmol, 4 eq.) was added dropwise. The reaction was then allowed to stir for 20 min at RT. After 20 min, the reaction had turned dark grey. The vial was transported to a cold bath at 0 °C and allowed to equilibrate for 10 min. After equilibration, allyl bromide (45 *u*L, 0.5 mmol, 2 eq.) was added. The reaction was then allowed to stir for another 20 min and finally acetophenone was added. The reaction was then stirred vigorously for 20 h and slowly turned light gray or greenish gray. The reaction was quenched by slow addition of a saturated sodium bicarbonate solution (approx. 1 mL). The mixture was separated and the aqueous layer washed with diethyl ether (1x1ml). The organic layers were combined and passed over a Celite plug and collected. The organic layer was then dried with sodium sulfate and passed over a silica plug. The silica was

rinsed with diethyl ether (~1mL). Samples of the corresponding enantiomeric products were then prepped for analysis without consideration for yield.

(S) 2-phenylpent-4-en-2-ol. Known Compound.[7] Enantiomeric analysis for (S) 2-phenylpent-4-en-2-ol was carried using a Thar SFC system. Concentrations of approximately 5-10 mg/mL sample were loaded. Injections were made at 5 uL sample loading on a Chiralcel AD-H (0.46cm x 25 cm) column. The chiral separation was carried out using 3% IPA at 40 °C. Peak 1 R_T: 10.1 min, Peak 2 R_T: 11.4.

Allylation of Methyl Ethyl Ketone

To an oven dried 1.5 dram vial with stir bar, was added $CrCl_3(THF)_3$ (9.4 mg, 0.025mmol, 0.1 eq.), 325 mesh manganese powder (27.5 mg, 0.5 mmol, 2 eq.), and ligand (.0275 mmol, 0.11 eq.). The vial was then fitted with a Tefton septum and cap. The vial was further sealed by wrapping the cap with Teflon tape. The vial was then purged under vacuum and flushed with argon gas 3 times. Under an atmosphere of argon, a standard solution of THF/TEA (0.6% TEA/THF, or 7.5 *u*L per 1.25 ml, in the test reaction 1.25 ml of this solution was to each vial) was added. The reaction was then stirred for 3-5 min and TMSCl (124 *u*L, 1 mmol, 4 eq.) was added dropwise. The reaction was then allowed to stir for 20 min at RT. After 20 min, the reaction had turned dark grey. Allyl bromide (45 *u*L, 0.5 mmol, 2 eq.) was added. The reaction was then allowed to stir for another 20 min and finally methyl ethyl ketone was added. The reaction was then stirred vigorously for 20h and slowly turned light gray or greenish gray. The reaction was quenched by slow addition of a saturated sodium bicarbonate solution (approx. 1 mL). The mixture was separated and the aqueous layer washed with diethyl ether (1x1ml). The organic layers were combined and passed over a Celite plug and collected. The organic layer was then dried with sodium sulfate and passed over a silica plug. The silica was rinsed with diethyl ether

(~1mL). Samples of the corresponding enantiomeric products were then prepped for analysis without consideration for yield.

3-methylhex-5-en-3-ol. Known Compound.[28] 3-methylhex-5-en-3-ol was separated by gas chromatography with a chiral stationary phase. A J&W Scientific 19091G-B213 HP-Chiral-20B cylcodex-B column was used. The length was 30m and I.D. was 0.320 mm. The run was isocratic at 45 °C and 0.6 mL/min flow rate for a 1 ul injection. The average retention time was 33.5 min and 34.5 min.

Data and Statistical Analysis

A detailed step by step summary of how the model for the 3×3 prediction of the best catalyst for benzaldehyde the raw data for which is given in Table 2.3. The data set consists of the nine ligands listed and each replicate run is included individually to account for experimental error. To develop a linear least squares model first we must define a few matrices. The design matrix is the matrix where each row represents an individual ligand is shown in Table 2.4. The design matrix was created by taking each individual term and performing the mathematical function described at the top of the column. The column for X^2 is simply the adjusted Charton value of X for an individual ligand squared, for YX^2 it is the value for Y multiplied by the value for X^2 It is important that the design matrix values match up with their corresponding $\Delta\Delta G^{\ddagger}$ values in the response matrix, as we are essentially solving a system of equations. A response matrix is also created which consists of the measured enantiomeric ratios given as $\Delta\Delta G^{\ddagger}$ values. The response matrix is given the designation Y and is the $\Delta\Delta G^{\ddagger}$ column in Table 2.3. The design matrix given the designation X and the two are manipulated according to the Equation 2.6. The resulting matrix C is the solution matrix of coefficient values z0-j in Equation 2.4. These values are listed in Table 2.5 which can be substituted into the original equation. The next step would

Table 2.3. Raw data for the 3×3 library used to determine Equation 2.11.

Substituent		Adj. Charton Values X	Y	$\Delta\Delta G^{\ddagger}$
X_{H}	Y_{Me}	-0.62	-0.93	0.022
X_{H}	Y_{tBu}	-0.62	-0.21	0.985
X_{H}	Y_{CEt3}	-0.62	0.93	0.153
X_{Me}	Y_{Me}	-0.1	-0.93	0.131
X_{Me}	Y_{tBu}	-0.1	-0.21	0.900
X_{Me}	Y_{CEt3}	-0.1	0.93	0.220
X_{tBu}	Y_{Me}	0.62	-0.93	0.230
X_{tBu}	Y_{tBu}	0.62	-0.21	0.360
X_{tBu}	Y_{CEt3}	0.62	0.93	0.022
X_{H}	Y_{Me}	-0.62	-0.93	0.022
X_{H}	Y_{tBu}	-0.62	-0.21	0.900
X_{H}	Y_{CEt3}	-0.62	0.93	0.180
X_{Me}	Y_{Me}	-0.1	-0.93	0.131
X_{Me}	Y_{tBu}	-0.1	-0.21	0.823
X_{Me}	Y_{CEt3}	-0.1	0.93	0.197
X_{tBu}	Y_{Me}	0.62	-0.93	0.243
X_{tBu}	Y_{tBu}	0.62	-0.21	0.312
X_{tBu}	Y_{CEt3}	0.62	0.93	0.022

Table 2.4. Initial design matrix for the allylation of benzaldehyde.

		Design Matrix X									
Substituent		**z0**	**X**	**Y**	**X^2**	**Y^2**	**XY**	**X^3**	**Y^3**	**YX^2**	**XY^2**
X_H	Y_{Me}	1	-0.62	-0.93	0.38	0.86	0.58	-0.24	-0.80	-0.36	-0.54
X_H	Y_{tBu}	1	-0.62	-0.21	0.38	0.04	0.13	-0.24	-0.01	-0.08	-0.03
X_H	Y_{CEt3}	1	-0.62	0.93	0.38	0.86	-0.58	-0.24	0.80	0.36	-0.54
X_{Me}	Y_{Me}	1	-0.1	-0.93	0.01	0.86	0.09	0.00	-0.80	-0.01	-0.09
X_{Me}	Y_{tBu}	1	-0.1	-0.21	0.01	0.04	0.02	0.00	-0.01	0.00	0.00
X_{Me}	Y_{CEt3}	1	-0.1	0.93	0.01	0.86	-0.09	0.00	0.80	0.01	-0.09
X_{tBu}	Y_{Me}	1	0.62	-0.93	0.38	0.86	-0.58	0.24	-0.80	-0.36	0.54
X_{tBu}	Y_{tBu}	1	0.62	-0.21	0.38	0.04	-0.13	0.24	-0.01	-0.08	0.03
X_{tBu}	Y_{CEt3}	1	0.62	0.93	0.38	0.86	0.58	0.24	0.80	0.36	0.54
X_H	Y_{Me}	1	-0.62	-0.93	0.38	0.86	0.58	-0.24	-0.80	-0.36	-0.54
X_H	Y_{tBu}	1	-0.62	-0.21	0.38	0.04	0.13	-0.24	-0.01	-0.08	-0.03
X_H	Y_{CEt3}	1	-0.62	0.93	0.38	0.86	-0.58	-0.24	0.80	0.36	-0.54
X_{Me}	Y_{Me}	1	-0.1	-0.93	0.01	0.86	0.09	0.00	-0.80	-0.01	-0.09
X_{Me}	Y_{tBu}	1	-0.1	-0.21	0.01	0.04	0.02	0.00	-0.01	0.00	0.00
X_{Me}	Y_{CEt3}	1	-0.1	0.93	0.01	0.86	-0.09	0.00	0.80	0.01	-0.09
X_{tBu}	Y_{Me}	1	0.62	-0.93	0.38	0.86	-0.58	0.24	-0.80	-0.36	0.54
X_{tBu}	Y_{tBu}	1	0.62	-0.21	0.38	0.04	-0.13	0.24	-0.01	-0.08	0.03
X_{tBu}	Y_{CEt3}	1	0.62	0.93	0.38	0.86	0.58	0.24	0.80	0.36	0.54

Table 2.5. Initial coefficient matrix.

Coefficient Matrix C
z0 = 0.757
a = -1.105
b = 0.280
c = -0.153
d = -0.720
f = -0.156
g = 0.444
h = -0.039
i = -0.096
j = 0.682

be the calculation of each of the errors for each of the terms. Examination of the errors reveals which terms to eliminate. The variance and covariance matrix (V) for these coefficients can be calculated by the Equation 2.7 and are shown in Table 2.6. The diagonal values highlighted along this matrix represent the nonscaled errors associated with each coefficient value. This clearly indicates that the gX^3 term should be omitted due its high error relative to the value of g given in Equation 2.14. Omission of this term leads to a new design matrix X_1 shown in Table 2.7. The coefficient values were calculated for this matrix along with the same response matrix Y to give matrix C_1 shown in Table 2.8. The corresponding variance-covariance matrix can be determined V_1 shown in Table 2.9. Table 2.9 indicates that the hY^3 term has the largest relative error and should be eliminated. Elimination of this term leads to design matrix X_2 (Table 2.10) with corresponding coefficient matrix C2 and variance-covariance matrix V2 (Tables 2.11 and 2.12 respectively). These results indicate that the iYX2 should be removed in the subsequent iteration.

The next iteration is shown in Tables 2.13, 2.14, and 2.15 which indicate the bY term should be removed from the model yielding the final design matrix shown in Table 2.16. This gives the final model presented in Equation 2.11 and coefficients given in Table 2.17. The final variance-covariance matrix is shown in Table 2.18 and demonstrates reasonable errors for coefficient. Removal of the term with the highest error cX^2 leads to greater errors for each term. Subsequent addition of each omitted term also led to greater overall error in the coefficients.

All other models were generated using this method. Analysis of variance was performed on each model and the relevant statistics are given in Table 2.19 where SS_{pe} is the sum of squares due to experimental error, SS_{lof} is the sum of squares due to lack of fit, S_r^2 is the sum of squares of the residuals and R^2 is the correlation coefficient and f is the Fischer statistic.

Table 2.6. Initial variance-covariance matrix.

	z0	*a*	*b*	*c*	*d*	*f*	*g*	*h*	*i*	*j*
z0	0.035	0.135	0.105	-0.037	-0.026	0	-0.349	-.123	0	0
a	0.135	1.350	-0.018	-0.365	0.004	0	-3.488	0.021	0	0
b	0.106	-0.018	0.518	0.005	-0.125	0	0.046	-0.602	0	0
c	-0.037	-0.365	0.005	0.099	-0.001	0	0.943	-0.006	0	0
d	-0.026	0.004	-.125	-0.001	0.030	0	-0.011	0.145	0	0
f	0	0	0	0	0	0	0	0	0	0
g	-0.349	-3.488	0.046	0.943	-0.122	0	9.011	-0.054	0	0
h	-0.123	0.020	-0.602	-0.006	0.145	0	-0.054	0.699	0	0
i	0	0	0	0	0	0	0	0	0	0
j	0	0	0	0	0	0	0	0	0	0

Table 2.7. First generation design matrix.

		Design Matrix X_1								
Substituent		**z0**	**X**	**Y**	**X^2**	**Y^2**	**XY**	**Y^3**	**YX^2**	**XY^2**
X_H	Y_{Me}	1	-0.62	-0.93	0.38	0.86	0.58	-0.80	-0.36	-0.54
X_H	Y_{tBu}	1	-0.62	-0.21	0.38	0.04	0.13	-0.01	-0.08	-0.03
X_H	Y_{CEt3}	1	-0.62	0.93	0.38	0.86	-0.58	0.80	0.36	-0.54
X_{Me}	Y_{Me}	1	-0.1	-0.93	0.01	0.86	0.09	-0.80	-0.01	-0.09
X_{Me}	Y_{tBu}	1	-0.1	-0.21	0.01	0.04	0.02	-0.01	0.00	0.00
X_{Me}	Y_{CEt3}	1	-0.1	0.93	0.01	0.86	-0.09	0.80	0.01	-0.09
X_{tBu}	Y_{Me}	1	0.62	-0.93	0.38	0.86	-0.58	-0.80	-0.36	0.54
X_{tBu}	Y_{tBu}	1	0.62	-0.21	0.38	0.04	-0.13	-0.01	-0.08	0.03
X_{tBu}	Y_{CEt3}	1	0.62	0.93	0.38	0.86	0.58	0.80	0.36	0.54
X_H	Y_{Me}	1	-0.62	-0.93	0.38	0.86	0.58	-0.80	-0.36	-0.54
X_H	Y_{tBu}	1	-0.62	-0.21	0.38	0.04	0.13	-0.01	-0.08	-0.03
X_H	Y_{CEt3}	1	-0.62	0.93	0.38	0.86	-0.58	0.80	0.36	-0.54
X_{Me}	Y_{Me}	1	-0.1	-0.93	0.01	0.86	0.09	-0.80	-0.01	-0.09
X_{Me}	Y_{tBu}	1	-0.1	-0.21	0.01	0.04	0.02	-0.01	0.00	0.00
X_{Me}	Y_{CEt3}	1	-0.1	0.93	0.01	0.86	-0.09	0.80	0.01	-0.09
X_{tBu}	Y_{Me}	1	0.62	-0.93	0.38	0.86	-0.58	-0.80	-0.36	0.54
X_{tBu}	Y_{tBu}	1	0.62	-0.21	0.38	0.04	-0.13	-0.01	-0.08	0.03
X_{tBu}	Y_{CEt3}	1	0.62	0.93	0.38	0.86	0.58	0.80	0.36	0.54

Table 2.8. First generation coefficient matrix.

Coefficient Matrix C_1
0.803
-0.561
0.012
-0.273
-0.746
-0.155
-0.070
-0.095
0.681

Table 2.9. First generation variance-covariance matrix

	z0	*a*	*b*	*c*	*d*	*f*	*h*	*i*	*j*
z0	0.224	0	1.077	0	-0.260	0	-1.252	0	0
a	0	0	0	0	0	0	0	0	0
b	1.077	0	5.177	0	-1.252	0	-6.018	0	0
c	0	0	0	0	0	0	0	0	0
d	-0.261	0	-1.252	0	0.303	0	1.455	0	0
f	0	0	0	0	0	0	0	0	0
h	-1.252	0	-6.018	0	1.455	0	6.996	0	0
i	0	0	0	0	0	0	0	0	0
j	0	0	0	0	0	0	0	0	0

Table 2.10. Second iteration of the design matrix.

		Design Matrix X_2							
Substituent		**z0**	**X**	**Y**	**X^2**	**Y^2**	**XY**	**YX^2**	**XY^2**
X_H	Y_{Me}	1	-0.62	-0.93	0.38	0.86	0.58	-0.36	-0.54
X_H	Y_{tBu}	1	-0.62	-0.21	0.38	0.04	0.13	-0.08	-0.03
X_H	Y_{CEt3}	1	-0.62	0.93	0.38	0.86	-0.58	0.36	-0.54
X_{Me}	Y_{Me}	1	-0.1	-0.93	0.01	0.86	0.09	-0.01	-0.09
X_{Me}	Y_{tBu}	1	-0.1	-0.21	0.01	0.04	0.02	0.00	0.00
X_{Me}	Y_{CEt3}	1	-0.1	0.93	0.01	0.86	-0.09	0.01	-0.09
X_{tBu}	Y_{Me}	1	0.62	-0.93	0.38	0.86	-0.58	-0.36	0.54
X_{tBu}	Y_{tBu}	1	0.62	-0.21	0.38	0.04	-0.13	-0.08	0.03
X_{tBu}	Y_{CEt3}	1	0.62	0.93	0.38	0.86	0.58	0.36	0.54
X_H	Y_{Me}	1	-0.62	-0.93	0.38	0.86	0.58	-0.36	-0.54
X_H	Y_{tBu}	1	-0.62	-0.21	0.38	0.04	0.13	-0.08	-0.03
X_H	Y_{CEt3}	1	-0.62	0.93	0.38	0.86	-0.58	0.36	-0.54
X_{Me}	Y_{Me}	1	-0.1	-0.93	0.01	0.86	0.09	-0.01	-0.09
X_{Me}	Y_{tBu}	1	-0.1	-0.21	0.01	0.04	0.02	0.00	0.00
X_{Me}	Y_{CEt3}	1	-0.1	0.93	0.01	0.86	-0.09	0.01	-0.09
X_{tBu}	Y_{Me}	1	0.62	-0.93	0.38	0.86	-0.58	-0.36	0.54
X_{tBu}	Y_{tBu}	1	0.62	-0.21	0.38	0.04	-0.13	-0.08	0.03
X_{tBu}	Y_{CEt3}	1	0.62	0.93	0.38	0.86	0.58	0.36	0.54

Table 2.11. Second iteration of the coefficient matrix.

Coefficient Matrix C_2
0.793
-0.562
0.021
-0.273
-0.686
-0.156
-0.096
0.682

Table 2.12. Second iteration of variance-covariance matrix.

	z0	***a***	***b***	***c***	***d***	***f***	***i***	***j***
z0	0.313	0.041	0.037	-0.479	-0.225	0.005	-0.061	-0.031
a	0.041	0.743	0.004	-0.058	-0.032	0.082	-0.004	-0.872
b	0.037	0.004	0.306	-0.057	-0.026	0.036	-0.797	-0.002
c	-0.479	-0.058	-0.057	1.858	0.000	-0.007	0.219	-0.001
d	-0.225	-0.032	-0.026	0.000	0.380	-0.004	0.006	0.053
f	0.005	0.082	0.036	-0.007	-0.004	0.372	-0.094	-0.094
i	-0.061	-0.004	-0.797	0.219	0.006	-0.094	3.033	-0.004
j	-0.031	-0.872	-0.002	-0.001	0.053	-0.094	-0.004	1.451

Table 2.13. Third iteration of the design matrix.

		Design Matrix X_3						
Substituent		**z0**	**X**	**Y**	**X^2**	**Y^2**	**XY**	**XY^2**
X_H	Y_{Me}	1	-0.62	-0.93	0.38	0.86	0.58	-0.54
X_H	Y_{tBu}	1	-0.62	-0.21	0.38	0.04	0.13	-0.03
X_H	Y_{CEt3}	1	-0.62	0.93	0.38	0.86	-0.58	-0.54
X_{Me}	Y_{Me}	1	-0.1	-0.93	0.01	0.86	0.09	-0.09
X_{Me}	Y_{tBu}	1	-0.1	-0.21	0.01	0.04	0.02	0.00
X_{Me}	Y_{CEt3}	1	-0.1	0.93	0.01	0.86	-0.09	-0.09
X_{tBu}	Y_{Me}	1	0.62	-0.93	0.38	0.86	-0.58	0.54
X_{tBu}	Y_{tBu}	1	0.62	-0.21	0.38	0.04	-0.13	0.03
X_{tBu}	Y_{CEt3}	1	0.62	0.93	0.38	0.86	0.58	0.54
X_H	Y_{Me}	1	-0.62	-0.93	0.38	0.86	0.58	-0.54
X_H	Y_{tBu}	1	-0.62	-0.21	0.38	0.04	0.13	-0.03
X_H	Y_{CEt3}	1	-0.62	0.93	0.38	0.86	-0.58	-0.54
X_{Me}	Y_{Me}	1	-0.1	-0.93	0.01	0.86	0.09	-0.09
X_{Me}	Y_{tBu}	1	-0.1	-0.21	0.01	0.04	0.02	0.00
X_{Me}	Y_{CEt3}	1	-0.1	0.93	0.01	0.86	-0.09	-0.09
X_{tBu}	Y_{Me}	1	0.62	-0.93	0.38	0.86	-0.58	0.54
X_{tBu}	Y_{tBu}	1	0.62	-0.21	0.38	0.04	-0.13	0.03
X_{tBu}	Y_{CEt3}	1	0.62	0.93	0.38	0.86	0.58	0.54

Table 2.14. Third iteration of the coefficient matrix.

Coefficient Matrix C_3
0.790
-0.562
-0.003
-0.266
-0.685
-0.158
0.681

Table 2.15. Third iteration of the variance-covariance matrix

	z0	*a*	*b*	*c*	*d*	*f*	*j*
z0	0.311	0.041	-0.475	0.041	-0.225	0.003	-0.031
a	0.041	0.743	-0.058	0.743	-0.032	0.082	-0.872
b	0.021	0.003	1.842	0.003	-0.025	0.012	-0.003
c	-0.475	-0.058	0.000	-0.058	0.000	0.000	-0.001
d	-0.225	-0.032	0.000	-0.032	0.380	-0.003	0.053
f	0.003	0.082	0.000	0.082	-0.003	0.369	-0.094
j	-0.031	-0.872	-0.001	-0.872	0.053	-0.094	1.451

Table 2.16. Final design matrix.

		Design Matrix X_4					
Substituent		**z0**	**X**	**X^2**	**Y^2**	**XY**	**XY^2**
X_H	Y_{Me}	1	-0.62	0.38	0.86	0.58	-0.54
X_H	Y_{tBu}	1	-0.62	0.38	0.04	0.13	-0.03
X_H	Y_{CEt3}	1	-0.62	0.38	0.86	-0.58	-0.54
X_{Me}	Y_{Me}	1	-0.1	0.01	0.86	0.09	-0.09
X_{Me}	Y_{tBu}	1	-0.1	0.01	0.04	0.02	0.00
X_{Me}	Y_{CEt3}	1	-0.1	0.01	0.86	-0.09	-0.09
X_{tBu}	Y_{Me}	1	0.62	0.38	0.86	-0.58	0.54
X_{tBu}	Y_{tBu}	1	0.62	0.38	0.04	-0.13	0.03
X_{tBu}	Y_{CEt3}	1	0.62	0.38	0.86	0.58	0.54
X_H	Y_{Me}	1	-0.62	0.38	0.86	0.58	-0.54
X_H	Y_{tBu}	1	-0.62	0.38	0.04	0.13	-0.03
X_H	Y_{CEt3}	1	-0.62	0.38	0.86	-0.58	-0.54
X_{Me}	Y_{Me}	1	-0.1	0.01	0.86	0.09	-0.09
X_{Me}	Y_{tBu}	1	-0.1	0.01	0.04	0.02	0.00
X_{Me}	Y_{CEt3}	1	-0.1	0.01	0.86	-0.09	-0.09
X_{tBu}	Y_{Me}	1	0.62	0.38	0.86	-0.58	0.54
X_{tBu}	Y_{tBu}	1	0.62	0.38	0.04	-0.13	0.03
X_{tBu}	Y_{CEt3}	1	0.62	0.38	0.86	0.58	0.54

Table 2.17. Final iteration of the coefficient matrix.

Coefficient Matrix C_4
0.92
-0.53
-0.89
-0.89
-0.69
-0.97

Table 2.18. Final iteration of the variance-covariance matrix.

	z0	***a***	***c***	***d***	***f***	***j***
z0	0.003	0.000	-0.005	-0.002	0.000	0.000
a	0.000	0.007	-0.001	0.000	0.001	-0.009
c	-0.005	-0.001	0.018	0.000	0.000	0.000
d	-0.002	0.000	0.000	0.004	0.000	0.001
f	0.000	0.001	0.000	0.000	0.004	-0.001
j	0.000	-0.009	0.000	0.001	-0.001	0.015

Table 2.19. Key statistics for each model.

	3x3			***5 x5***
	Benzaldehyde	***Acetophenone***	***Methyl Ethyl Ketone***	***Benzaldehyde***
SS_{pe}	0.008	.029	0.001	0.029
SS_{lof}	0.065	0.986	0.986	0.872
S_r^2	0.006	.084	0.008	0.021
R^2	0.98	0.91	0.98	0.73
f	62	25	131	12
Confidence Level	95%	95%	95%	90%

References

(1) Jacobsen, E. N.; Pfaltz, A.; Yamamoto, H.; Editors *Comprehensive Asymmetric Catalysis I-III, Volume 2*; Springer, 1999.

(2) Jacobsen, E. N.; Pfaltz, A.; Yamamoto, H.; Editors *Comprehensive Asymmetric Catalysis I-III, Volume 1*; Springer, 1999.

(3) Jacobsen, E. N.; Pfaltz, A.; Yamamoto, H.; Editors *Comprehensive Asymmetric Catalysis I-III, Volume 3*; Springer, 1999.

(4) Walsh, P. J.; Kozlowski, M. C. *Fundamentals of Asymmetric Catalysis*; University Science Books: Sausalito, 2009.

(5) Miller, J. J.; Sigman, M. S. *Angew. Chem., Int. Ed.* **2008**, *47*, 771.

(6) Harper, K. C.; Sigman, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 2179.

(7) Miller, J. J.; Sigman, M. S. *J. Am. Chem. Soc.* **2007**, *129*, 2752.

(8) Miller, J. J.; Rajaram, S.; Pfaffenroth, C.; Sigman, M. S. *Tetrahedron* **2009**, *65*, 3110.

(9) Anderson, G. W.; Zimmerman, J. E.; Callahan, F. M. *J. Am. Chem. Soc.* **1967**, *89*, 5012.

(10) Han, S.-Y.; Kim, Y.-A. *Tetrahedron* **2004**, *60*, 2447.

(11) Phillips, A. J.; Uto, Y.; Wipf, P.; Reno, M. J.; Williams, D. R. *Org. Lett.* **2000**, *2*, 1165.

(12) Charton, M. *J. Org. Chem.* **1976**, *41*, 2217.

(13) Charton, M. *J. Am. Chem. Soc.* **1975**, *97*, 3691.

(14) Charton, M. *J. Am. Chem. Soc.* **1975**, *97*, 3694.

(15) Charton, M. *J. Am. Chem. Soc.* **1975**, *97*, 1552.

(16) Sigman, M. S.; Miller, J. J. *J. Org. Chem.* **2009**, *74*, 7633.

(17) Stewart, F. M. *Introduction to linear algebra*; Van Nostrand: Princeton, N.J., 1963.

(18) Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*; University Science Books: Sausalito, 2006.

(19) Jones, J. E. *Proc. R. Soc. A* **1924**, *106*, 441.

(20) Jones, J. E. *Proc. R. Soc. A* **1924**, *106*, 463.

(21) Deming, S. N.; Morgan, S. L. *Experimental Design: A Chemometric Approach. Second, Revised and Expanded Edition*; Elsevier, 1993.

(22) Hansch, C.; Leo, A. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, DC, 1995.

(23) Denmark, S. E.; Gould, N. D.; Wolf, L. M. *J. Org. Chem.* **2011**, *76*, 4337.

(24) Brown, J. M.; Deeth, R. J. *Angew. Chem. Int. Ed.* **2009**, *48*, 4476.

(25) Donoghue, P. J.; Helquist, P.; Norrby, P.-O.; Wiest, O. *J. Am. Chem. Soc.* **2008**, *131*, 410.

(26) Weill, N.; Corbeil, C. R.; De Schutter, J. W.; Moitessier, N. *J. Comput. Chem.* **2011**, *32*, 2878.

(27) Rajaram, S.; Sigman, M. S. *Org. Lett.* **2005**, *7*, 5473.

(28) Shi, S.-L.; Xu, L.-W.; Oisaki, K.; Kanai, M.; Shibasaki, M. *J. Am. Chem. Soc.* **2010**, *132*, 6638.

CHAPTER 3

3D FREE ENERGY RELATIONSHIPS AND THE PROPARGYLATION OF KETONES

Introduction

Our success in examining Nozaki-Hiyama-Kishi (NHK) reactions using the oxazoline-proline ligand library led us to examine the 9-membered library in several reactions we had explored previously without success.[1] We examined the crotylation, vinylation and propargylation of ketones (Figure 3.1). Among these reactions, we focused on the propargylation of ketones because of the potential to expand our correlations to include electronic effects in addition to steric effects in asymmetric catalysis.[2]

One of the first enantioselective catalytic propargylation of aldehydes was reported by Keck and coworkers in 1994 using catalytic Ti(IV) with a Binol ligand and allenyltin.[3] Since that time there have been several additional reports of enantioselective carbonyl propargylation. Interest in propargylation has been due to the rich chemistry available for alkyne manipulation. Propargylation products are often considered to be complimentary to allylation products but propargylation products offer an orthogonal array of reactions for synthetic chemistry. Ketone propargylation products can be considered a masked allylation product, which can be accessed through partial-reductions, or they can be considered as a completely different synthon. Ketone propargylation products have been vetted in synthetic chemistry for a number of different reactions including cross-coupling, cycloaddition, oxidation, halogenation, and

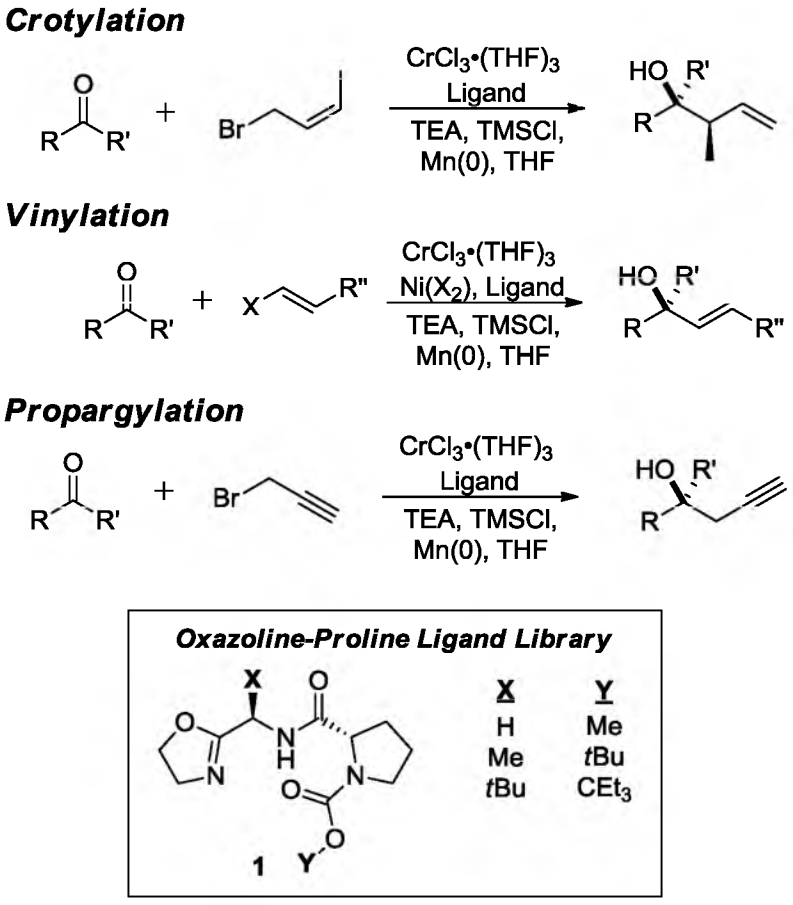


Figure 3.1. Challenging reactions evaluated using the oxazoline-proline ligand library.

metathesis. As the utility of propargylation products expanded, the propargylation of aldehydes was the predominant focus only recently were reports published on the propargylation of ketones.[4-7]

Until 2010, there were no reports of catalytic enantioselective ketone propargylation reactions of any type. In 2004, Soderquist and coworkers developed a series of chiral allenylboronates capable of addition to carbonyls.[4] Their report represented one of the few ways to access enantiomerically enriched ketone propargylation products until Shibasaki and coworkers reported the first enantioselective catalytic Cu(I)-mediated addition of allenylboronates to ketones (Figure 3.2A).[5] They used a complex modular amino-phosphine ligand to impart facial bias. They reported the successful propargylation of 14 different ketones, but found that aliphatic ketones suffered from lower enantioselectivity and presented a challenging substrate class. Later, Fandrick and coworkers improved this reaction considerably through the use of BINAP as the ligand and a protected propargylboronate, which broadened the scope of the reaction to include aliphatic ketones (Figure 3.2B).[6] Since these reports and our own report, Schaus and coworkers have demonstrated the enantioselective propargylation of ketones using organocatalytic bisphenols, again making use of allenylboronates (Figure 3.2C).[7] All of these examples employ allenyl- or propargylboronate, in part because it is an attractive propargyl fragment source. However, collectively their data indicate that the scope of substrates compatible with the allenylboronates is limited and the more compatible propargylboronates must be protected in order to achieve desired results.

In contrast to this limitation, propargylation under NHK conditions could allow for direct access of propargylation products through a propargylic halide. Prior to these reports, Kishi had shown that enantioselective propargylation of aldehydes was compatible with NHK conditions

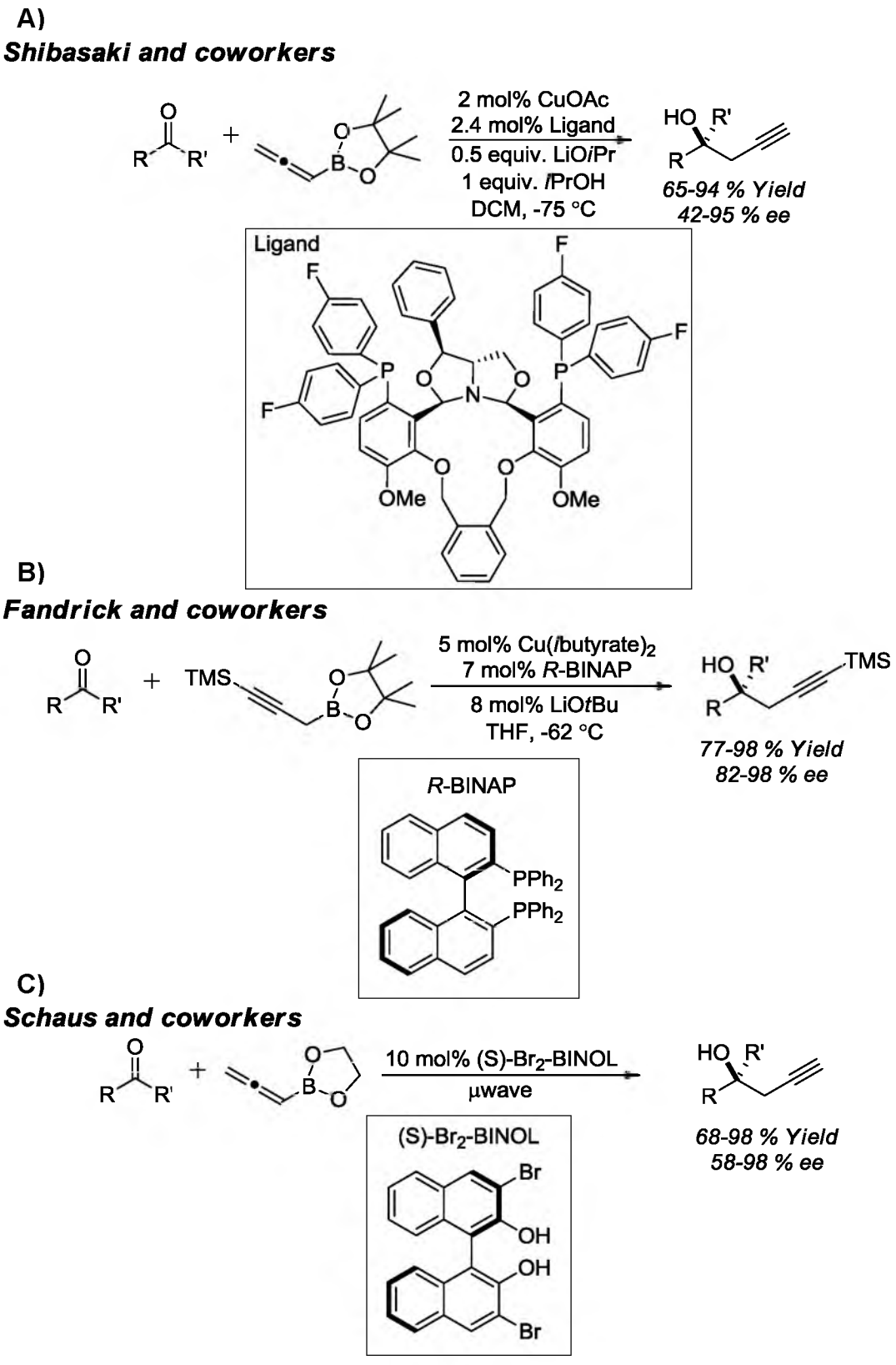


Figure 3.2. Recently reported asymmetric propargylation methodologies.

using propargyl halides.[8] Due to our group's previous success expanding the scope of NHK allylation reactions to include ketones, we decided to pursue the enantioselective propargylation of ketones.[9] Initially, Dr. Tejas Pathak evaluated the reaction by synthesizing several variants of the oxazoline-proline ligand scaffold and evaluating them for enantioselectivity. Under standard conditions, reaction with ketone substrates was achieved; however, enantioselectivity proved far more difficult to optimize and eventually the reaction was abandoned. After establishing our steric-based 3D LFER approach, we evaluated the reactions in Figure 3.1 including the propargylation of ketones.

Library Evaluation of the NHK Propargylation of Ketones

To reinitiate our study of ketone propargylation, we used the same conditions as with previous studies, substituting propargyl bromide for allyl bromide. The model substrate selected for evaluation of the oxazoline-proline library was acetophenone due to its simplicity (Figure 3.3A). Using the techniques and data analysis discussed in the previous chapter, the propargylation of acetophenone was evaluated using the 3×3 ligand library. Using adjusted Charton values, the data was fit to a 3rd order polynomial resulting in Equation 3.1 and the surface constructed shown in Figure 3.3B.

$$\Delta\Delta G^{\ddagger} = 0.30 - 0.48X + 0.19Y - 0.72Y^2 - 0.56XY + 0.39XY^2 - 0.98YX^2 \quad (3.1)$$

The model gives a domed surface, which has a maximum point nearest the X_{Me} Y_{tBu} ligand. This ligand had been experimentally evaluated and gave only 75:25 er. This quantified maximum point falls well below our threshold of acceptability, which prompted us to draw the conclusion that the oxazoline-proline ligand structure was not capable of inducing adequate enantioselectivity and a major ligand redesign would be required.

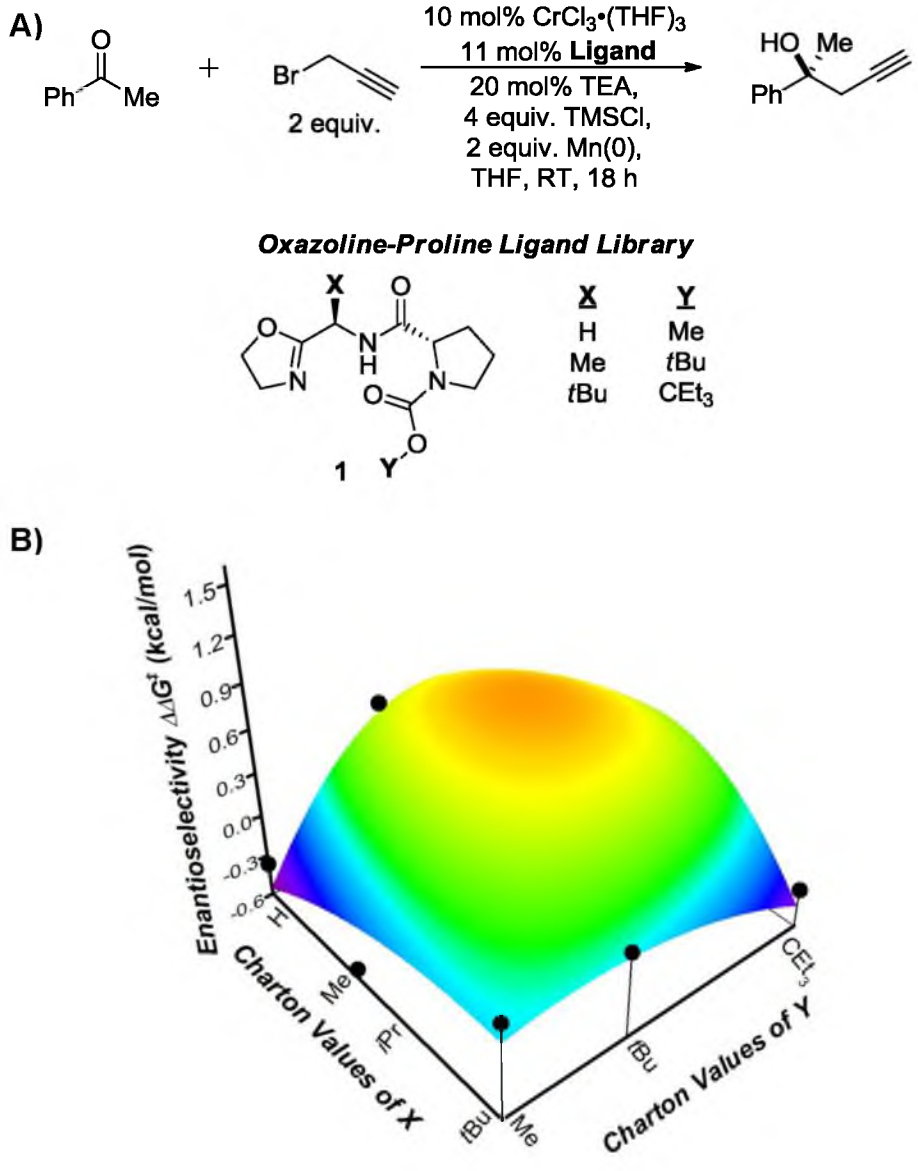


Figure 3.3. A) Evaluation of the propargylation of acetophenone using the oxazoline-proline ligand library. B) The surface described by Equation 1.

Ligand Redesign

Two key elements were incorporated into the ligand redesign and they were largely based on data that Dr. Tejas Pathak had gathered in his initial studies of the propargylation reaction. Our first key observation arose from our evaluation of the 3x3 library. From the oxazoline-proline library, ligand **1a** was found to be the highest performing in terms of enantioselectivity with 75:25 er (Figure 3.4). Dr. Pathak's data suggested that substitution on the oxazoline positively affected enantioselectivity. Ligands **2** and **3** demonstrate an apparent steric effect by achiral substituents on the oxazoline. However, the performance of ligands **4a** and **4b** indicate that an additional chiral center on the oxazoline does not affect the facial selectivity or the enantioselectivity. In fact, the results from **4a** and **4b** demonstrate that the addition of a chiral substituent on the oxazoline does not affect enantioselectivity, when compared to **4c** which lacks substitution at this position. Taken together, these results suggest that steric interactions in the plane of the oxazoline are capable of impacting enantioselectivity. The other key piece of data provided by Dr. Pathak was a study he performed probing the electronic nature of the substrate shown in Figure 3.5 where a significant electronic dependence on the substrate was observed. A Hammett plot constructed from the data revealed good correlation between enantioselectivity and substrate electronic nature.

These results shaped the redesign of the oxazoline-proline scaffold. We hypothesized that the electronic effect exhibited by the substrate might be mirrored in the catalyst. In order to examine catalyst electronic effects, the oxazoline would need to be substituted with a modular arene capable of electronic variation. Synthesis and evaluation of ligand **5** demonstrated the viability of pyridine as a substitute for the oxazoline (Figure 3.6). Ligand **5** gave only comparable results with oxazoline-proline ligands; hence, we hypothesized that the steric effect exhibited by ligands **2** and **3** might be mimicked by extending the pyridine along the

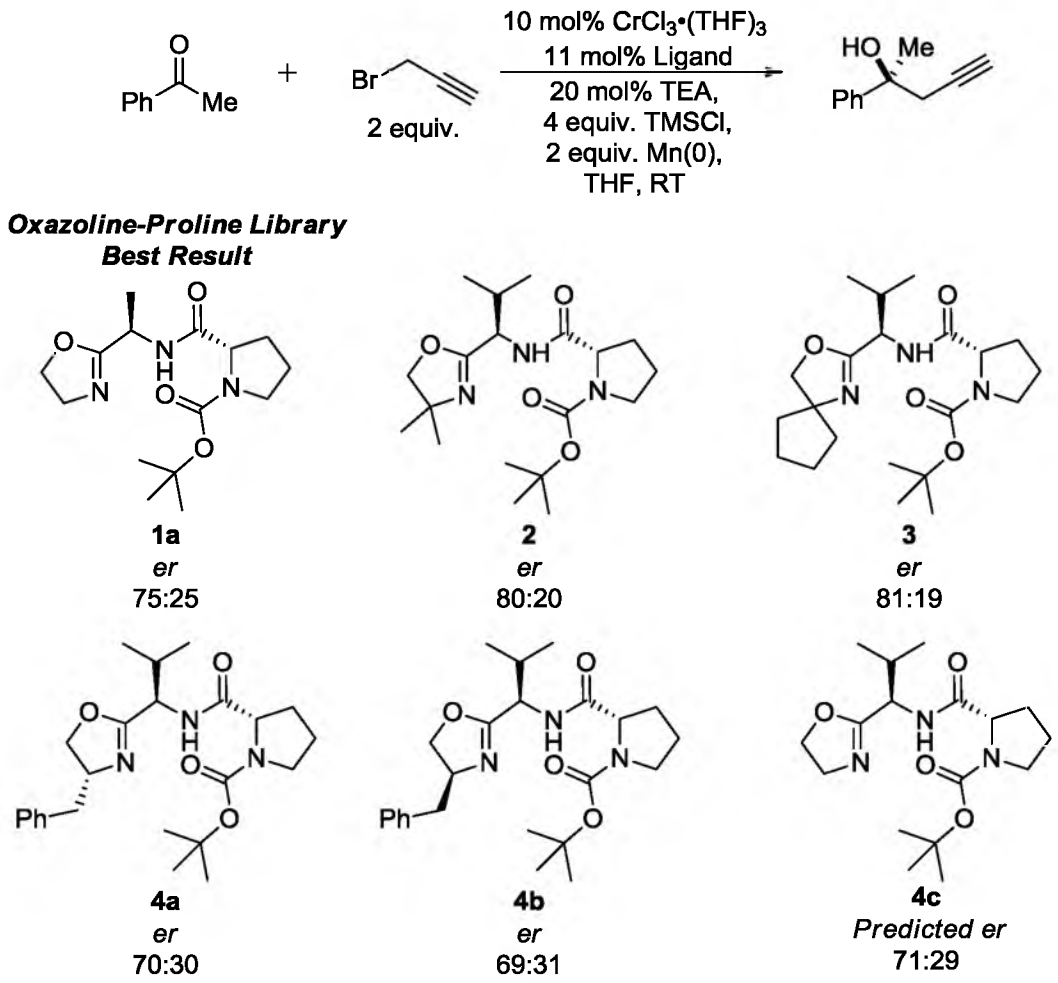
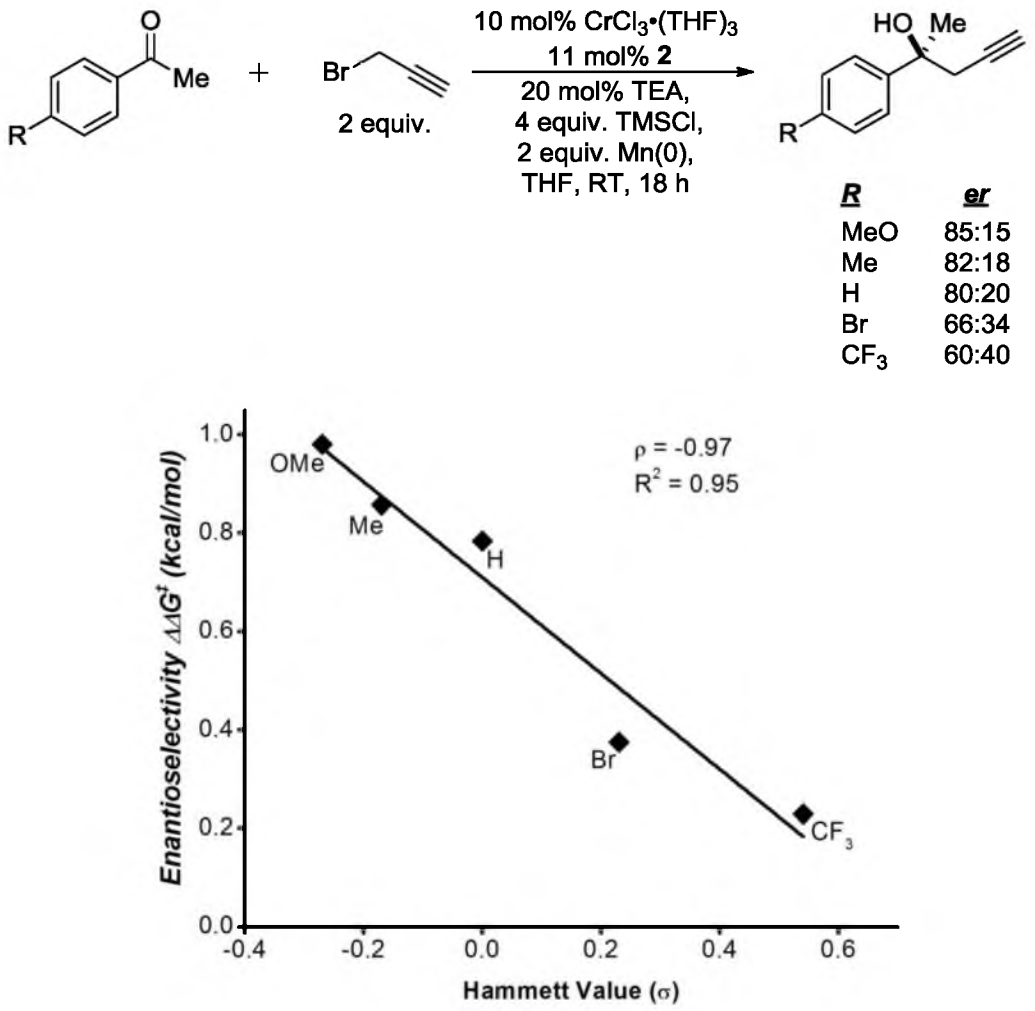
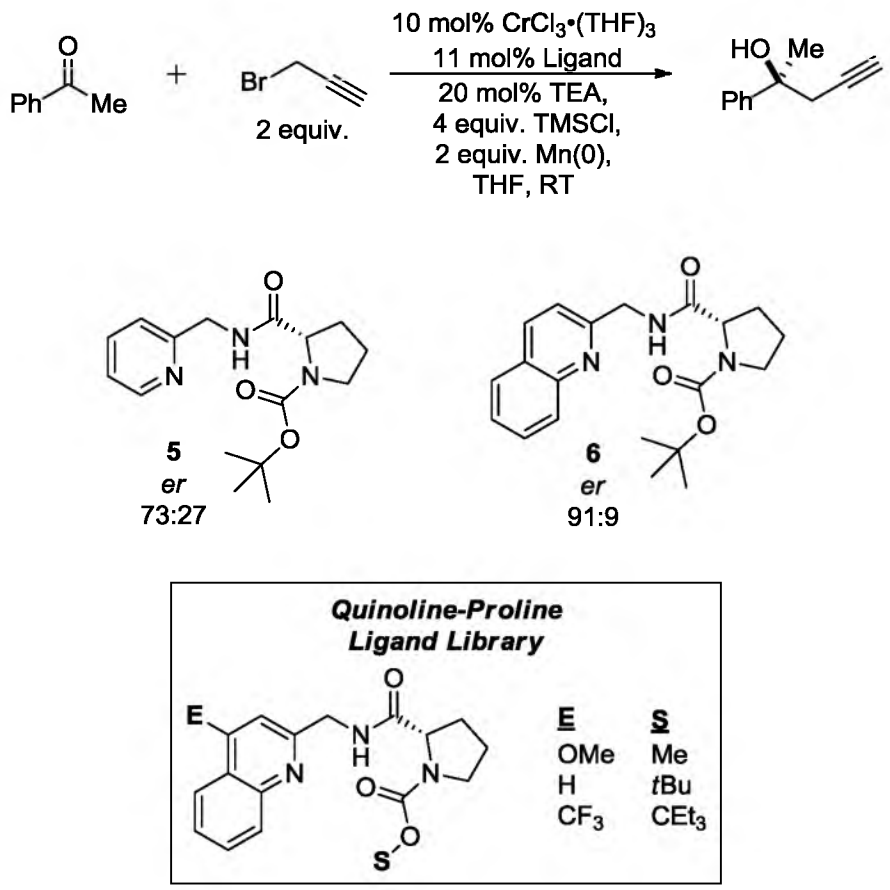


Figure 3.4. Key ligands whose results in the propargylation of acetophenone directed the redesign of the ligand scaffold.



Figures 3.5. Correlation of substrate electronics in the NHK propargylation of acetophenone.



Figures 3.6. Replacement of the oxazoline module with heteroaromatics leading to the development of a new ligand scaffold.

plane of the arene. Excitingly, evaluation of the quinoline-proline (QuinPro) ligand **6** afforded a significant increase in enantioselectivity to 91:9 er. This ligand framework possessed the necessary features to simultaneously examine electronic effects and steric effects on enantioselectivity. To study both effects simultaneously, a new library of ligands, termed QuinPro, was synthesized and evaluated for the propargylation of acetophenone.

In developing the QuinPro library, the principles of experimental design were used to arrive at a model that encapsulated the available steric substituents as well as electronic substituents. A 3×3 design was chosen again to balance predictive power with synthetic effort. To model the steric effect of the carbamoyl group (S), we examined the same substituents as the oxazoline-proline 3×3 library again following the rationale that these substituents encompass the available synthetic space for the Charton steric parameter. To probe the electronic substituent (E), we chose Hammett sigma values. Although these values are typically used to describe electronic perturbation of nonheterocyclic arenes, the assumption was made they would scale to describe similar resonance and inductive effects on the quinoline ring. Figure 3.7 shows our experimental design for the substituent variation for both the E and S positions.

Synthesis of the Quinoline-Proline Ligand Library

To examine the steric effect of the carbamoyl group, we relied on the same reaction sequences used to develop the oxazoline-proline ligand library. The larger CEt_3 substituent had to be installed by synthesis of carbonate **7** followed by condensation of **7** onto benzyl-protected proline (Scheme 3.1). Hydrogenolysis yielded the CEt_3 carbamate-protected proline. Synthesis of **8** was accomplished under standard conditions and commercially available Boc-proline was used for the S_{tBu} ligands.

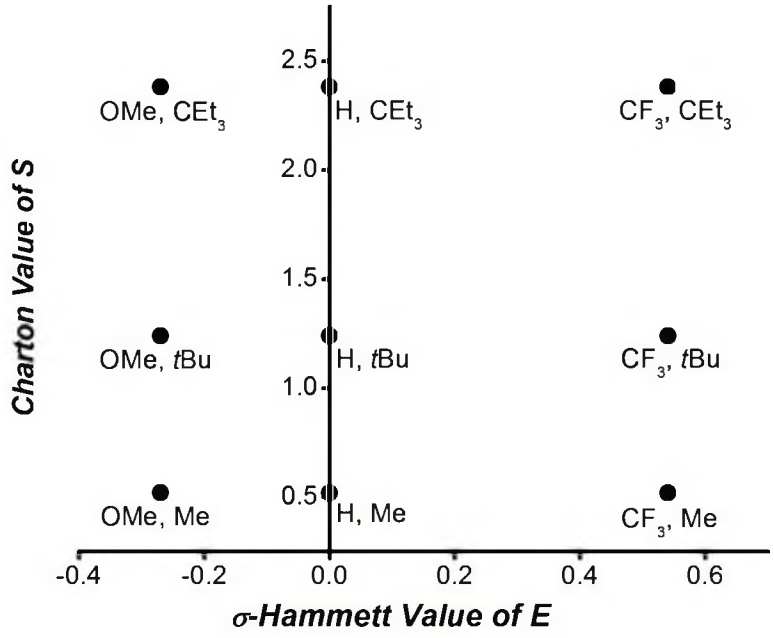
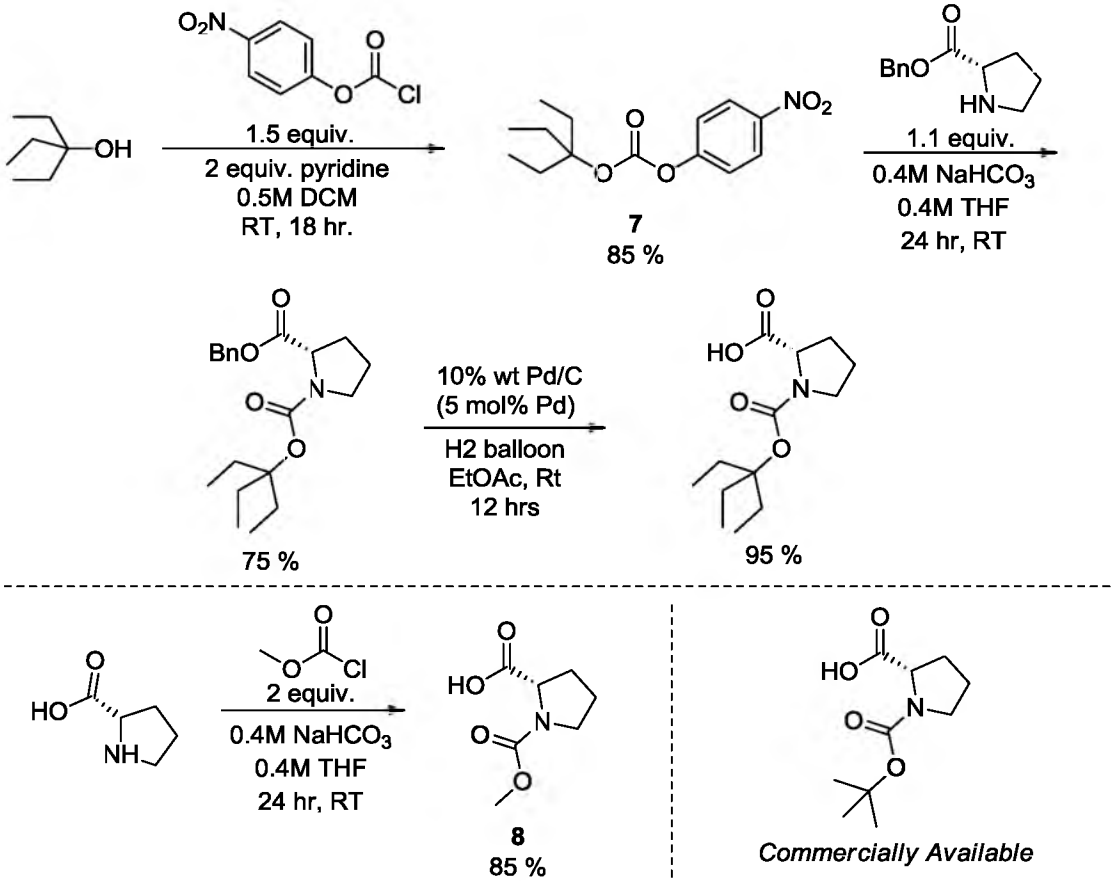


Figure 3.7. Layout of the steric and electronic variation in the experimental design for the propargylation of acetophenone.

Scheme 3.1. Synthesis of sterically modified prolines.

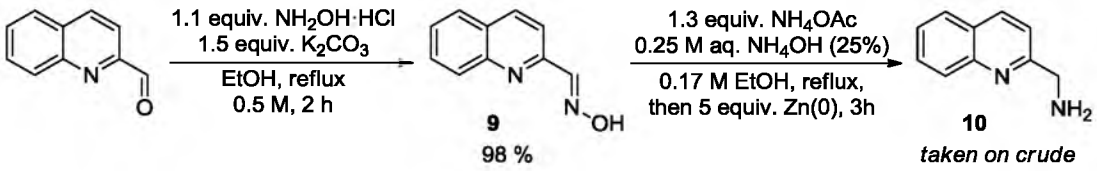


The synthesis of the oxazoline-proline library featured substituents that could be installed via the same synthetic route because variation of the amino acid partners had little effect on reaction outcomes. Synthesis of the quinoline rings proved much more difficult. Each quinoline substituent required different starting materials and different reaction pathways to arrive at the desired amines for the final coupling reaction. The most direct route begins with commercially available 2-quinoline-carboxaldehyde (Scheme 3.2). Condensation with hydroxylamine hydrochloride yields oxime **9** that can then be reduced using super-stoichiometric amounts of Zn in the presence of ammonium salts. The resulting amine was taken on crude to the final coupling reaction.

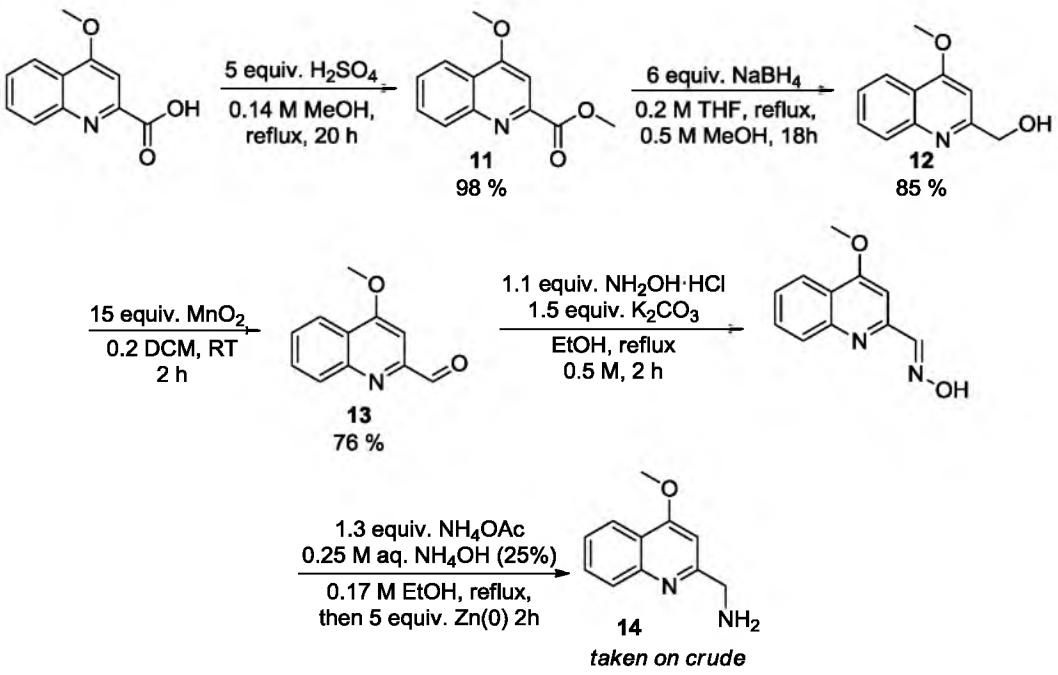
The synthesis of the 4-methoxy quinoline could be achieved only by incorporation of the methoxy group into the starting material (Scheme 3.3). From commercially available 4-methoxyquinoline-2-carboxylic acid, esterification under acidic conditions gave **11**, which proved difficult to reduce. Attempts at partial reduction of **11** using diisobutylaluminium hydride all gave the over-reduced product in low yield. Use of stronger reducing agents such as lithium aluminum hydride also gave lower yields, presumably due to the strong chelating product, which created emulsions during workup and prevented good isolation. A higher yielding reduction of **11** was finally discovered using sodium borohydride and THF/MeOH. Oxidation of **12** using manganese dioxide gave aldehyde **13,** which could undergo the previously described sequential reductive amination to **14**.

The 4-trifluoromethyl quinoline derivative proved the most difficult to access. Using Schlosser and coworkers' previously reported route from commercially available starting materials, we were able to isolate significant quantities of **15** (Scheme 3.4).[10] Carbonylation of

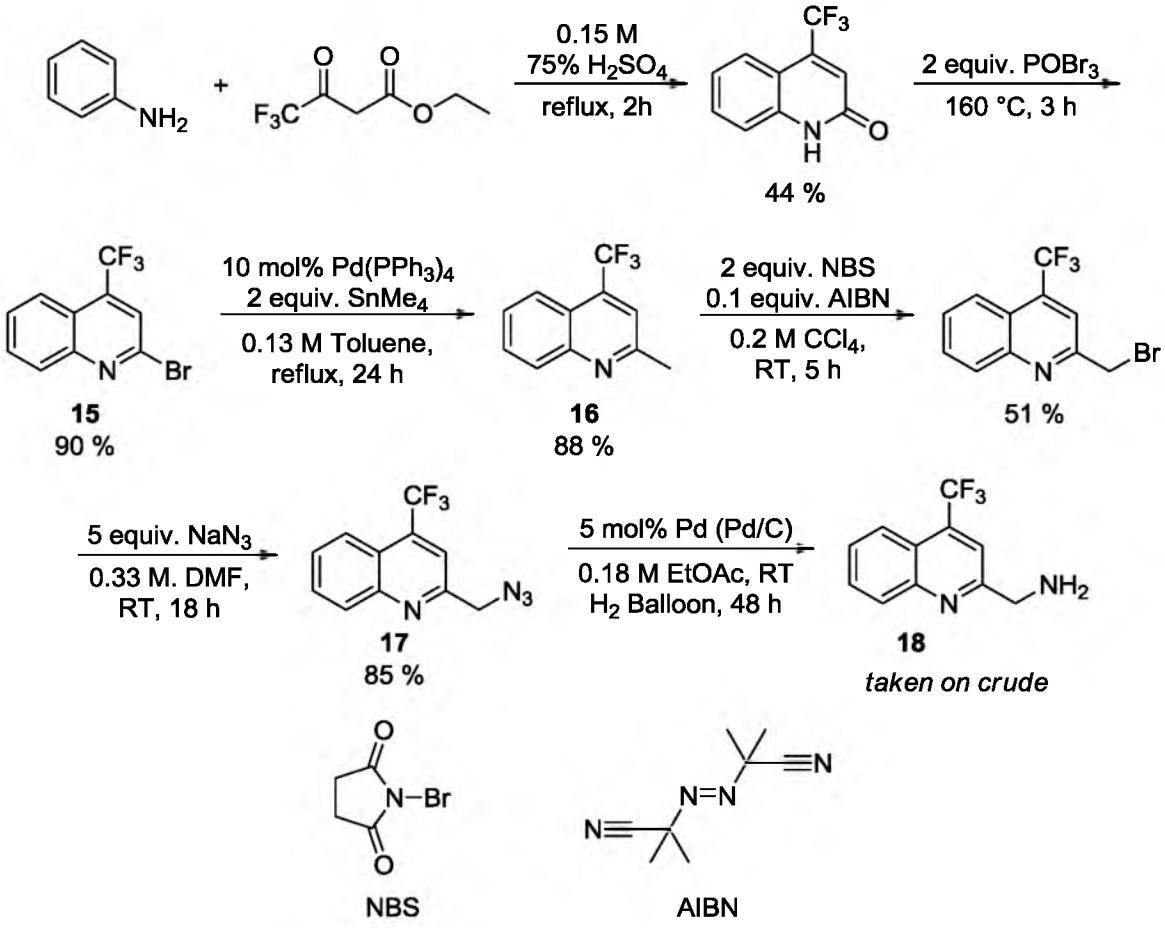
Scheme 3.2. Synthesis of 4-hydrogen quinolinemethylamine.



Scheme 3.3. Synthesis of 4-methoxyquinolinemethylamine.



Scheme 3.4. Synthesis of 4-trifluoroquinolinemethylamine



15 had been reported by Schlosser and coworkers via lithiation and quenching using CO_2. Numerous attempts at this reaction had been performed in our lab but we had been unable to reproduce Schlosser and coworkers' results. To overcome this problem, a Stille cross-coupling reaction between **15** and tetramethyltin was used to generate **16** in good yield. Treatment of **16** with NBS and AIBN gave moderate yields of single bromination product, due to a competitive dibromination pathway. S_N2 displacement of the bromide by azide gave **17** which could be reduced catalytically with palladium on carbon to give amine **18**.

Amines **10**, **14** and **18** all proved to be difficult to purify, handle and store. As such each was taken on crude to the final coupling reaction (Scheme 3.5). Using Anderson's conditions, the coupling reaction was low yielding presumably due to the use of crude amines as reagents, which only allowed the estimation of their equivalency. Following these synthetic schemes, the quinoline-proline library was synthesized and evaluated under the same conditions demonstrated in Figure 3.3A. Again, each data point was replicated over a minimum of two experiments and the raw data are given in Table 3.1 along with the parameter values used to develop the model.

Model Determination for the Quinoline-Proline Library

Once synthesis and evaluation was complete, the remaining task was to model the data. To do so, the full 3rd order polynomial was evaluated according to Equation 3.2 where E represents the electronic perturbation described by Hammett values and S describes the steric variation described by Charton values.

$$\Delta\Delta G^{\ddagger} = z0 + aE + bS + cE^2 + dS^2 + iES + gE^3 + hS^3 + iSE^2 + jES^2 \quad (3.2)$$

The treatment of the data employed similar techniques to those previously outlined. Accordingly, a model was derived to fit the data and shown in by Equation 3.3.

Scheme 3.5. Peptide coupling.

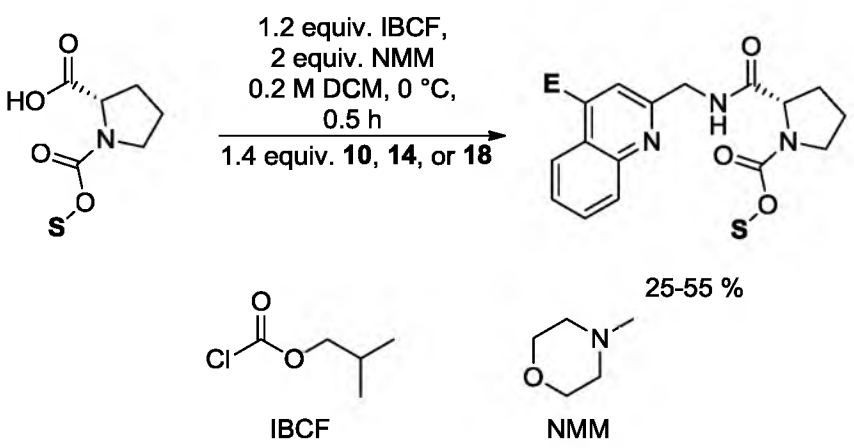


Table 3.1. Raw data for the propargylation of acetophenone.

E	S	*Hammett-σ Value*	*Charton Value*	*Average er*	*Average ΔΔG‡*
CF_3	CH_3	0.54	0.52	46:54	-0.10
CF_3	tBu	0.54	1.24	56:44	0.15
CF_3	CEt_3	0.54	2.38	61:39	0.26
H	CH_3	0	0.52	54:46	0.09
H	tBu	0	1.24	87:13	1.12
H	CEt_3	0	2.38	78:22	0.73
OCH_3	CH_3	-0.27	0.52	60:40	0.24
OCH_3	tBu	-0.27	1.24	91:9	1.39
OCH_3	CEt_3	-0.27	2.38	81:19	0.85

$$\Delta\Delta G^{\ddagger} = -1.20 + 1.22E + 2.84S - 0.85S^2 - 3.79ES + 1.25ES^2 \quad (3.3)$$

The surface described by Equation 3.3 is shown in Figure 3.8. The statistically significant model shows dependence on both the electronic and steric terms. Interestingly, two crossterms remain in the final model indicating that the steric and electronic elements are interacting synergistically. The model gives a high correlation coefficient ($R^2 = 0.95$) indicating a good fit of the data. Because all nine of the ligands that we had synthesized were incorporated into the training set, no independent validation could be explored. Therefore, in order to validate the model, a technique termed leave-one-out validation (LOO) was utilized. As implied, the technique omits a single data point from the training set. The remaining data points are used to calculate a new model based on the original. The new model can then be used to predict the value of the omitted point. The process is repeated until each data point has been omitted and predicted. The linear regression between the predicted values and experimentally observed values gives a Q^2 value of 0.89 for Equation 3.3 (Figure 3.9). Comparison of these this Q^2 with those developed in the previous chapter would be misleading. The validations performed in the previous chapter were external validations not LOO validations. Typically, LOO Q^2 values are typically higher because the model is reiterated for each missing data point. However, comparison reported LOO values indicates that these values are remarkably high and suggest a high predictive power.

The model revealed that the surface maximum lies near the E_{OMe} S_{tBu} ligand data point. The maximum lies closer to reported Charton values in the S dimension but because of some doubts we had in those values (explained in Chapter 4), and the economy of the E_{OMe} S_{tBu} ligand, we elected not to extrapolate the relationship further. It should also be noted that the E_{OMe} S_{tBu} ligand was included in the training set so this is not an a priori determination of optimal ligand

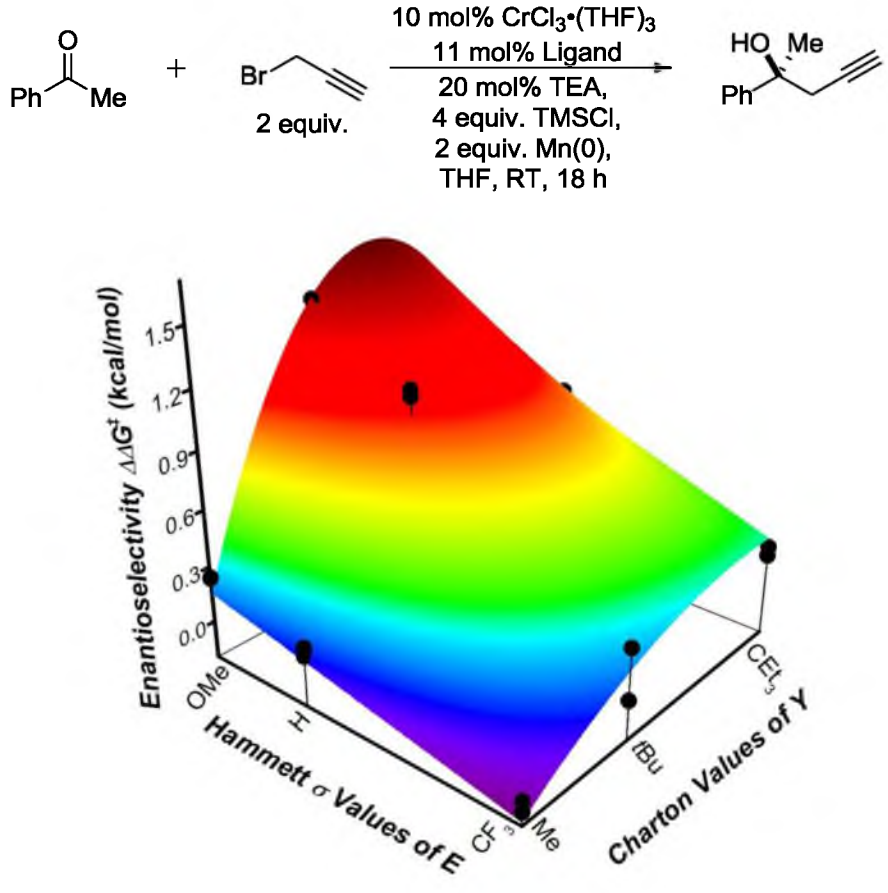


Figure 3.8. 3D LFER developed for the propargylation of acetophenone as described by Equation 3.

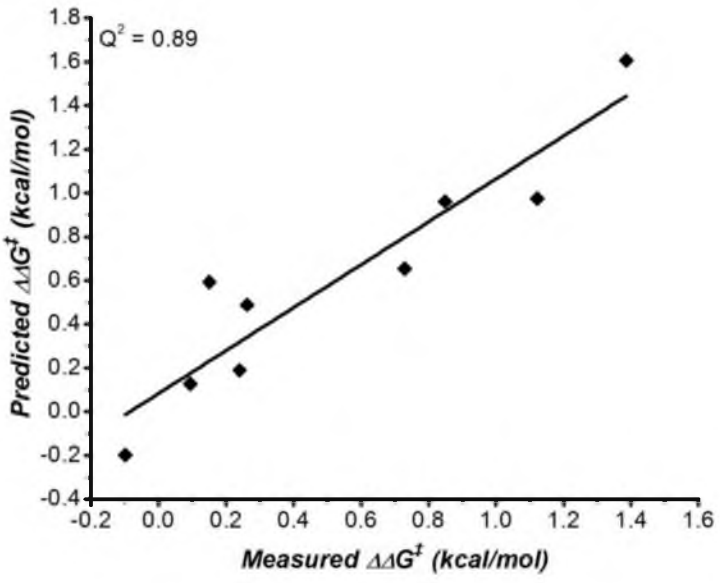


Figure 3.9. Results of the Leave-One-Out validation performed on Equation 3.

properties. However, the robustness of the model supports that any improvement in enantioselectivity that could be gained through interpolation of the maximum point would be marginal and not worth the extra synthetic effort required to generate the catalyst.

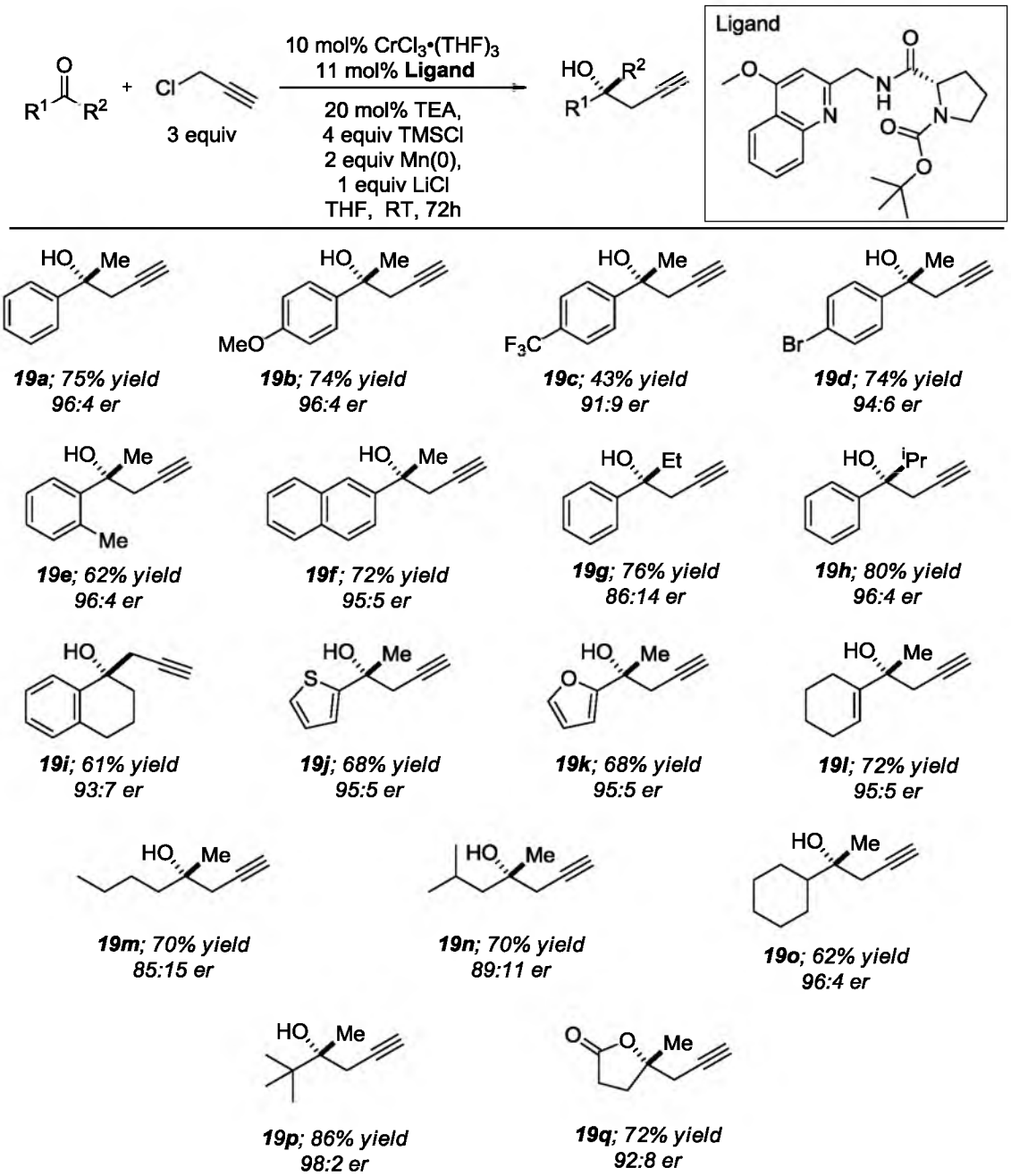
The standard approach to asymmetric catalyst development is to evaluate catalyst properties independently and individually. Inspection of the surface shown in Figure 3.8 reveals that the standard approach would be unlikely to expose the synergy between the steric and electronic effects in this catalyst system. For example, steric effects are substantially mitigated for electron poor ligands, whereas a considerable enhancement of the steric effect is observed for electron rich ligands. Furthermore, the electronic effect for smaller substituents is diminished as compared to the effect observed for the S_{tBu} series. Simply put, a poor choice of one variable could not be superseded by an ideal choice in the other.

This empirical inspection is strongly reinforced by the derived mathematical model described in Equation 3.3. Inclusion of crossterms between the putatively independent variables not only considerably improves the model but also mathematically defines a significant synergistic relationship between the electronic and steric interactions in the catalyst. Specifically, the terms of ES and ES^2 suggest that the electronic nature of the catalyst and the steric environment in the transition state are closely correlated. Assuming that the electronic variation is directly affecting the Lewis-acidity of the Cr-center, the importance of the crossterms on enantioselectivity can be hypothesized through invoking the Hammond postulate. The effect of an electron donating substituent would mitigate the Lewis-acidity of the Cr(III) catalyst. A less Lewis-acidic Cr-catalyst would require greater proximity to the substrate in order to activate it for nucleophilic attack. The result would be a later more product-like transition state.

A major concern in the optimization of an enantioselective catalytic reaction is over-optimizing catalyst properties to a specific substrate. Typically, the desired result of reaction optimization is a highly selective catalyst that is general over a broad substrate scope. The 3D-LFER approach to asymmetric catalysis has the potential to over-optimize a specific reaction. To combat this potential problem, the QuinPro library was simultaneously evaluated with an aliphatic substrate, namely 2-hexanone. These results reveal similar trends.

With some confidence in the optimal ligand structure, the reaction conditions were further optimized. The only notable change was the use of propargyl chloride, which gave a modest increase in enantioselectivity (93:7 er to 96:4 er for acetophenone). Otherwise, condition manipulation led to little or no change in enantioselectivity. The use of propargyl chloride led to a significant reduction in the reaction rate. To mitigate this rate reduction exhibited by propargyl chloride, several additives were explored in the reaction with lithium chloride proving most beneficial, as previously reported by Kishi.[11] With the optimized ligand and conditions, the scope of the reaction was examined. Generally, high enantioselectivity and good yield are observed for aryl methyl ketones as shown in Figure 3.10, with a modest drop in enantioselectivity for electron poor substrates. Other aryl ketones gave good to excellent enantiomeric ratios, as highlighted by substrate **19h**. Heteroaromatics are also well-tolerated under the reaction conditions, with thiophene and furan derived ketones undergoing highly enantioselective propargylation reactions (**19j** and **19k**). Additionally, an α,β-unsaturated ketone is an excellent substrate for enantioselective propargylation, leading to a 95:5 er (**19l**).

Excitingly, the catalyst was found to be competent in the allylation of aliphatic ketones, a challenging substrate class. Facial selectivity appears to be based on steric differentiation, giving higher enantiomeric ratios with increased steric bulk on a single side of the ketone (**19m-**



Yields are the average of at least two experiments.

Figure 3.10. The scope of the propargylation of ketones using the optimal ligand derived from the QuinPro library.

p). An er of 85:15 observed for 2-hexanone (**19m**) is impressive, considering it involves differentiating a methyl from an *n*-butyl group. Groups with substitution at the α-position of the ketone significantly enhance the enantiomeric ratios, as highlighted by ketones with a cyclohexyl (96:4 er, **19o**) and a t-butyl group (98:2 er, **19p**). Furthermore, a γ-butyrolactone (**19q**) can be synthesized with good enantiomeric ratio from propargylation of an aliphatic ketone with a pendant ester.

Conclusions

An exciting expansion of our 3D-QSAR technique was explored in the NHK propargylation of ketones. The initial evaluation of the 3×3 library guided our efforts away from the oxazoline-proline ligand scaffold. The strength of the model developed from the oxazoline-proline library gave us confidence to undertake a major redesign of the ligand. The redesign was hypothesis driven but supported by empirical observation. Eventually, we discovered a new ligand framework capable of imparting satisfactory enantioselectivity. The synthesis of the QuinPro library was arduous in comparison to the previous libraries, but allowed for determination of the optimal ligand substituents.

The catalyst determined by this approach proved to be give high enantioselectivities across a number of substrates. The development of this reaction gives a pattern of library evaluation followed by hypothesis driven redesign. This same pattern has been implemented in asymmetric catalysis for years, but has been driven by empirical results. Our pattern is driven primarily by multiple quantified relationships, which provide a picture of which elements are affecting selectivity and how they might be interacting.

Not only did the 3D-QSAR models quantifiably guide us to an optimal ligand, they also generated cross-relationships between isolated variables. The crossterms relating electronic

character to steric effects allowed us to generate multiple hypotheses about the character of the transition state that were not solely based on kinetic studies or spectroscopic data. The relationship itself was nonintuitive and has generated greater understanding in how enantioselectivity might be accomplished in the system.

Experimental

General Information

Unless otherwise noted, all reactions were performed under an argon atmosphere with stirring. Toluene, dichloromethane, dichloroethane, and THF were dried before use by passing through a column of activated alumina. Methanol was distilled from magnesium methoxide. Triethylamine was distilled from CaH_2. Dimethylformamide and dimethylacetamide were dried by leaving over molecular sieves for 72 h prior to use. Acetophenone was purified by drying over Na_2SO_4 then fractional distillation. Propargyl bromide was purchased as a solution of 80% in Toluene and used without further purification. Propargyl chloride was purified by fractional distillation and stored with protection from light. $CrCl_3(THF)_3$ was prepared by soxhlet extraction of anhydrous $CrCl_3$, in the presence of a minimal amount of zinc, with anhydrous THF. Ketone substrates present in the reaction scope table were purified by fractional distillation, passage through a column of alumina or recrystallization from ethanol and ether. All other reagents were used without further purification. Yields were calculated for material judged homogeneous by thin-layer chromatography and ^1H NMR. Thin-layer chromatography was performed with EMD silica gel 60 F254 plates eluting with the solvents indicated, visualized by a 254 nm UV lamp, and stained either with potassium permanganate, phosphomolybdic acid, or ninhydrin with charring. Flash column chromatography was performed with EcoChrom MP Silitech 32-63D 60Å silica gel, slurry packed with solvents indicated in glass columns. Nuclear

magnetic resonance spectra were acquired at 300, 400, or 500 MHz for ^{1}H, and 75, 100, or 125 MHz for ^{13}C. NMR spectra for ligands were collected at 50 °C for proline containing ligands and ligand precursors. Chemical shifts for proton nuclear magnetic resonance (^{1}H NMR) spectra are reported in parts per million relative to the line of $CHCl_3$ singlet referenced to 7.24 ppm. Chemical shifts for carbon nuclear magnetic resonance (^{13}C NMR) spectra are reported in parts per million downfield relative to the center-line of the $CDCl_3$ triplet at 77.23 ppm. The abbreviations s, d, t, p, sep, dd, td, bs, and m stand for the resonance multiplicities singlet, doublet, triplet, pentet, septet, doublet of doublets, triplet of doublets, broad singlet, and multiplet, respectively. Optical rotations were obtained (Na D line) using a Perkin Elmer Model 343 Polarimeter fitted with a micro cell with a 1 dm path length. Concentrations are reported in g/100 mL. SFC (super critical fluid chromatography) analysis was performed at 25 °C or 40 °C, using a Thar instrument fitted with chiral stationary phase (as indicated). Melting points were obtained on an electrothermal melting point apparatus and are uncorrected. Unless otherwise noted, glassware for all reactions was oven-dried at 110 °C and cooled in a dry atmosphere prior to use.

Ligand Synthesis and Characterization

Synthetic schemes are shown beginning from commercially available materials. Full characterization of previously unpublished compounds is contained herein.

Methyl 4-methoxyquinoline-2-carboxylate (11). 4-Methoxy-2-quinolinecarboxylic acid (5.100 g., 25.1 mmol, 1 equiv) was placed in dry 250 ml rb flask. Dry MeOH (125 mL, 5ml/mmol) was then added. To the slurry was added concentrated H_2SO_4 0.28 ml/mmol) and the reaction mixture was heated to reflux. After 24 h, the reaction mixture was cooled to RT and then slowly poured into 500 ml of saturated aqueous $NaHCO_3$. The mixture was then placed in an ice bath

for 1 h and the precipitate collected by filtration. The precipitate was then dried on a vacuum line overnight. 98% Yield (5.398 g). Colorless solid. R_f = 0.2 (50% Ace/Hex). MP = 142-143 °C. ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.13 (m, 2H), 7.71 (t, *J* = 8, 1H), 7.54 (m, 2H), 4.07 (s, 3H), 4.03 (s, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 166.4, 163.5, 149.2, 148.5, 130.7, 130.3, 127.8, 122.2, 121.9, 100.3, 56.3, 53.4. HRMS $C_{12}H_{11}NO_3$ $(M+Na)^+$ calcd. 240.0643, obsvd. 240.0631.

(4-Methoxyquinolin-2-yl)methanol (12). A slurry of methyl 4-methoxyquinoline-2-carboxylate (5.000 g, 23.04 mmol, 1 equiv) was created by addition of dry THF (115 ml, 5 ml/mmol) in a 500 mL rb flask. The flask was fitted a condenser and heated to reflux. Finely powdered $NaBH_4$ (5.230 g, 138.2 mmol, 6 equiv) was added down the condenser slowly over a period of 1 h. Then dry MeOH (46 mL, 2 ml/mmol) was added down the condenser slowly over a period of 1 h. The reaction mixture was then heated overnight at reflux. Upon completion as confirmed by TLC after 26 h, the reaction mixture was cooled to RT and saturated aqueous NH_4Cl (100 mL) was added slowly by addition funnel. The reaction mixture was then concentrated in vacuo to remove all THF and extracted with DCM (3x100 mL). The organic extracts were combined and dried with Na_2SO_4 and concentrated. The material was then purified via flash chromatography on silica using 50% Acetone/Hexanes. 85% Yield (3.705 grams). R_f = 0.1 (35% Ace/Hex). MP = 102-103 °C. ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.12 (d, *J* = 8, 1H), 7.95 (d, *J* = 8, 1H), 7.66 (d, *J* = 8, 1H), 7.47 (d, *J* = 8, 1H), 6.58 (s, 1H), 4.82 (s, 2H), 4.4 (bs, 1H), 3.99 (s, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 163.00, 160.60, 147.81, 130.24, 128.31, 125.62, 122.09, 121.12, 96.97, 64.64, 55.94. HRMS $C_{11}H_{12}NO_2$ $(M+H)^+$ calcd. 190.0863, obsvd. 190.0871.

4-methoxyquinoline-2-carbaldehyde (13). In a round bottom flask was placed (4-Methoxyquinolin-2-yl)methanol (2.500 g, 13.4 mmol, 1 equiv.) which was dissolved in dry DCM (67 mL, 5 mL/mmol) and a stirbar added. MnO_2 (17.5 g, 201 mmol, 15 equiv.) was added in

portions over 5 min The reaction was then allowed to stir overnight. Upon completion by TLC, The heterogeneous mixture was filtered through celite, the filter cake was washed by ~500 mL of MeOH and all of the organic filtrate was concentrated an purified by column chromatography on silica using 10% EtOAc/Hexanes. 76% Yield (1.901 grams). R_f = 0.1 (10% EtOAc/Hex). MP = 120-121 °C. ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 9.12 (s, 1H) 8.12 (d, *J* = 8, 1H), 7.95 (d, *J* = 8, 1H), 7.66 (d, *J* = 8, 1H), 7.47 (d, *J* = 8, 1H), 6.58 (s, 1H), 3.99 (s, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 188.23 163.23, 160.54, 147.83, 131.55, 129.24, 123.24, 122.97, 121.82, 99.43, 55.40. HRMS $C_{11}H_{12}NO_2$ $(M+H)^+$ calcd. 190.0863, obsvd. 190.0871.

4-methoxyquinoline-2-carbaldehyde oxime. 4-methoxyquinoline-2-carbaldehyde (995 mg, 1 equiv., 5.32 mmol) was dissolved in 95% EtOH (26 mL, 5 mL/mmol) in a round bottom flask with stirrer. Hydroxylamine hydrochloride (443 mg, 1.2 equiv., 6.38 mmol) was dissolved in water (5.3 mL, 1 mL/mmol) in a separate vial and K_2CO_3 (956 mg, 1.3 equiv., 6.98 mmol) was added and effervescence observed. This aqueous solution was then added to the EtOH containing the starting material and the flask fitted with a condenser and refluxed for 90 min. The crude mixture was then evaporated under reduced pressure until a solid remained. The solid was dissolved in acetone (100 mL) and filtered. The fitrate was evaporated yielding a colorless solid. This material was taken on without further purification.

(4-Methoxyquinolin-2-yl)methanamine (14). The 4-methoxyquinoline-2-carbaldehyde (1 equiv. 5.32 mmol) oxime was dissolved in EtOH (30 mL, 6 mL/mmol) in a round bottom flask with stirrer and NH_4OAc (532 mg, 1.3 equiv., 6.9 mmol) was added followed by a solution of 25% NH4OH (20 mL, 3.66 mL/mmol) in H_2O. The flask was then fitted with a condenser and heated to reflux. Once at reflux, powdered zinc (1.739 g, 5 equiv., 26.6 mmol) was added down the condenser in small portions over a period of 2 h. Once all the zinc was added, the reaction mixture was allowed to reflux for an additional h and then cooled to room temperature. Crude

material was then stirred with a saturated solution of KOH (25 mL) in water and DCM (60 mL). After 30 min of stirring, the layers were allowed to resolve and the organic layer was isolated and dried with Na_2SO_4. It was then evaporated under reduced pressure to yield a yellow oil, which was taken on in the aforementioned amide coupling reaction without considerations for purification.

General Ligand Coupling Procedure

In a separate dry round bottom flask, the appropriate proline derivative (1 equiv) was dissolved in dry DCM (4 ml/mmol). NMM (1.5 equiv) was added and the reaction mixture cooled in an ice bath. After 10 min, IBCF (1 equiv) was added dropwise and the reaction mixture stirred in the ice bath for 30 min and a colorless precipitate was observed. Then, (4-Methoxyquinolin-2-yl)methanamine was added in a minimal amount of dry DCM by syringe dropwise. The reaction mixture was generally observed to turn green at this point. After 4 h, the reaction mixture was quenched by dilution in DCM (10 times dilution) and washed with saturated aqueous $NaHCO_3$ (10-40 mL) and H_2O (10-40 mL). The organic layer was then dried with Na_2SO_4 and concentrated. The final ligand was purified by sequential columns eluted in mixtures of acetone/hexanes and methanol/dichloromethane to assume maximum purity.

(S)-Methyl 2-((4-methoxyquinolin-2-yl)methylcarbamoyl)pyrrolidine-1-carboxylate. See general ligand coupling procedure. Colorless Solid. $[\alpha]^{20}_{D} = -30.6°$ (c = 0.354, $CHCl_3$). $R_f = 0.1$ (50% Ace/Hex). Purified in 50% Acetone/Hexane. 45% Yield. NMRs were obtained at higher temperature due to rotamers. ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.14 (d, *J* = 8.4, 1H), 7.93 (d, *J* = 8.4, 1H), 7.65 (t, *J* = 8.0, 1H), 7.45 (t, *J* = 8.0, 1H), 6.68 (s, 1H) 4.65 (d, *J* = 4.8, 2H), 4.30 (m, 1H), 3.75-3.65 (m, 8H), 2.2-1.8 (m, 4H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 172.1, 156.5, 130.2, 128.7, 125.7, 122.1, 121.0, 120.4, 98.6, 68.1, 61.3, 56.0, 52.9, 48.93, 29.8, 25.9. Did not produce both

carbonyl carbons even at extended relaxation time high numbers of scans. HRMS $C_{18}H_{21}N_3O_4$ $(M+H)^+$ calcd. 343.1562, obsvd. 343.1569.

(S)-Tert-butyl 2-((4-methoxyquinolin-2-yl)methylcarbamoyl)pyrrolidine-1-carboxylate. See general ligand coupling procedure. Purified via sequential columns in acetone/hexanes and methanol/dichloromethane. 62% Yield. Yellowish Solid. $[\alpha]^{20}_D$ = -27.1° (c = 0.524, $CHCl_3$). R_f = 0.3 (50% Ace/Hex). NMRs were obtained at higher temperature due to rotamers, which persisted. ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.09 (m, 1H), 7.93 (m, 1H), 7.65 (m, 1H), 7.43 (m, 1H), 6.92 (d, *J* = 8, 1H), 6.68 (s, 1H), 5.01 (m, 1H), 4.80 (m, 1H), 4.64 (d, *J* = 2.4, 1H), 4.00 (m, 3H), 3.75-3.4 (m, 2H), 2.3-1.8 (m, 4H), 1.44 (m, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 163.1, 158.7, 148.93, 148.6, 129.9, 129.7, 128.6, 125.6, 125.5, 125.3, 121.9, 121.9, 121.8, 99.3, 98.7, 80.3, 79.3, 61.4, 56.9, 55.9, 54.4, 53.6, 47.4, 47.3, 45.5, 30.5, 28.7, 28.5, 28.3, 24.6. HRMS $C_{21}H_{28}N_3O_4$ $(M+H)^+$ calcd. 386.2074, obsvd. 386.2091.

(S)-3-Ethylpentan-3-yl 2-((4-methoxyquinolin-2-yl)methylcarbamoyl)pyrrolidine-1-carboxylate. See general ligand coupling procedure. Colorless Solid. $[\alpha]^{20}_D$ = -24.3° (c = 0.427, $CHCl_3$). R_f = 0.4 (50% Ace/Hex). 32% Yield. NMRs were obtained at higher temperature due to rotamers,which persisted. ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.11 (m, 1H), 7.98 (m, 1H), 7.69 (m, 1H), 7.45 (m, 1H), 6.93 (d, *J* = 8, 1H), 6.78 (s, 1H), 5.12 (m, 1H), 4.95 (m, 1H), 4.69 (d, *J* = 2.4, 1H), 4.03 (s, 3H), 3.75-3.4 (m, 2H), 2.3-1.6 (m, 10H), 0.70 (s, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 163.1, 158.7, 148.93, 148.6, 129.9, 129.7, 128.6, 125.6, 125.5, 125.3, 121.9, 121.9, 121.8, 99.3, 98.7, 80.3, 79.3, 61.4, 56.9, 55.9, 54.4, 53.6, 47.4, 47.3, 45.5, 30.5, 28.7, 28.5, 28.3, 24.6, 8.4. Did not produce both carbonyl carbons even at extended relaxation time high numbers of scans. HRMS $C_{24}H_{34}N_3O_4$ $(M+H)^+$ calcd. 428.2549, obsvd. 428.2552.

Quinolin-2-ylmethanamine (10). 2-quinolinecarboxaldehyde (630 mg, 1 equiv., 4 mmol) was dissolved in 95% EtOH (20 mL, 5 mL/mmol) in a round bottom flask with stirrer.

Hydroxylamine hydrochloride (278 mg, 1 equiv., 4 mmol) was dissolved in water (4 mL, 1 mL/mmol) in a separate vial and K_2CO_3 (663 mg, 1.2 equiv., 4.8 mmol) was added and effervescence observed. This aqueous solution was then added to EtOH and the flask fitted with a condenser and refluxed for 90 min. The crude mixture was then evaporated under reduced pressure until an orange solid remained. The solid was dissolved in DCM (30 mL) and washed with water (2x20 mL). The organic layer was then collected and dried with Na_2SO_4 and evaporated under reduced pressure again producing an orange solid. The material was taken on without further purification. Next, the solid was dissolved in EtOH (24 mL, 6 mL/mmol) in a round bottom flask with stirrer and NH_4OAc (400 mg, 1.3 equiv., 5.2 mmol) was added followed by a solution of 25% NH_4OH (14.5 mL, 3.66 mL/mmol) in H_2O. The flask was then fitted with a condenser and heated to reflux. Once at reflux, powdered zinc (1.308 g, 5 equiv., 20 mmol) was added down the condenser in small portions over a period of 2 h. Once all the zinc was added, the reaction mixture was allowed to reflux for an additional 1 h and then cooled to room temperature. Crude material was then stirred with a saturated solution of KOH (25 mL) in water and DCM (60 mL). After 30 min of stirring, the layers were allowed to resolve and the organic layer was isolated and dried with Na_2SO_4. It was then evaporated under reduced pressure to yield a thick reddish oil, which was taken on in the aforementioned amide coupling reaction without considerations for purification.

(S)-Methyl 2-(quinolin-2-ylmethylcarbamoyl)pyrrolidine-1-carboxylate. See general ligand coupling procedure. Yellow Oil. $[\alpha]^{20}_D$ = -64.9° (c = 1.768, $CHCl_3$). R_f = 0.2 (50% Acetone/Hex). Purified in 50% Acetone/Hexane. 28% Yield. NMRs were obtained at higher temperature due to rotamers. ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.04 (d, *J* = 4, 1H), 7.97 (d, *J* = 4, 1H), 7.73 (d, *J* = 4, 1H), 7.64 (t, *J* = 8, 1H), 7.46 (t, *J* = 8, 1H), 7.27 (d, *J* = 4, 1H), 4.68 (d, *J* = 2.6, 2H) 4.39 (bs, 1H), 3.65-3.48 (m, 2H), 2.22-1.83 (m, 4H) ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 163.1,

158.7, 147.6, 136.9, 129.82, 129.1, 127.7, 127.5, 126.5, 119.9, 61.2, 52.86, 47.4, 44.9, 29.6, 29.5, 24.5. Did not observe both carbonyl carbons even at extended relaxation time/high numbers of scans. HRMS $C_{17}H_{20}N_3O_3$ $(M+H)^+$ calcd. 314.1499, obsvd. 314.1513.

(S)-Tert-butyl-2-(quinolin-2-ylmethylcarbamoyl)pyrrolidine-1-carboxylate. See general ligand coupling procedure. Colorless Solid. $[\alpha]^{20}_D$ = -59.7° (c = 4.504, $CHCl_3$). R_f = 0.3 (50% Ace/Hex). Purified in 50% Acetone/Hexane. 21% Yield. NMRs were obtained at higher temperature due to rotamers. . ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.01 (d, *J* = 4, 1H), 7.96 (d, *J* = 4, 1H), 7.70 (d, *J* = 16, 1H), 7.61 (t, *J* = 8, 1H), 7.42 (t, *J* = 8, 1H), 7.26 (d, *J* = 4, 1H), 4.66 (d, *J* = 2.4, 2H) 4.31 (bs, 1H), 3.49-3.43 (m, 2H), 2.22-1.75 (m, 4H), 1.33 (s, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 172.8, 156.74, 14706, 136.9, 130.9, 129.8, 129.1, 127.7, 127.5, 126.5, 119.9, 80.3, 61.2, 47.2, 45.1, 31.1, 29.4, 28.5, 24.3. Did not observe both carbonyl carbons even at extended relaxation time/high numbers of scans. HRMS $C_{20}H_{26}N_3O_3$ $(M+Na)^+$ calcd. 378.1788, obsvd. 378.1792.

(S)-3-Ethylpentan-3-yl 2-(quinolin-2-ylmethylcarbamoyl)pyrrolidine-1-carboxylate. See general ligand coupling procedure. Colorless Solid. $[\alpha]^{20}_D$ = -53.6° (c = 1.586, $CHCl_3$). R_f = 0.4 (50% Ace/Hex). Purified in 30% Acetone/Hexane. 44% Yield. NMRs were obtained at higher temperature due to rotamers. ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.06 (d, *J* = 4, 1H), 8.00 (d, *J* = 4, 1H), 7.75 (d, *J* = 4, 1H), 7.66 (t, *J* = 8, 1H), 7.48 (t, *J* = 8, 1H), 7.31 (d, *J* = 4, 1H), 4.70 (m, 2H) 4.35 (bs, 1H), 3.54-3.44 (m, 2H), 2.35-1.6 (m, 10H), 0.69 (s, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 172.8, 156.7, 147.5, 136.9, 130.2, 129.7, 129.2, 127.7, 127.5, 126.5, 120.0, 88.1, 61.5, 47.3, 45.2, 31.4, 27.4, 24.0, 7.8. Did not produce both carbonyl carbons even at extended relaxation time high numbers of scans. HRMS $C_{24}H_{34}N_3O_4$ $(M+H)^+$ calcd. 428.2549, obsvd. 428.2552.

4-(Trifluoromethyl)quinolin-2(1H)-one. Known Compound.[10]

2-Bromo-4-(trifluoromethyl)quinoline (15). Known Compound.[10]

2-Methyl-4-(trifluoromethyl)quinoline (16). In a dry Schlenk bomb flask were added 2-bromo-4-(trifluoromethyl)quinoline (1.6g, 5.79 mmol, 1 equiv), tetrakistriphenylphosphinepalladium(0) (670 mg, .58 mmol, 0.1 equiv) and dry dimethylacetamide (30 mL, 5 ml/mmol). To this, tetramethyltin(IV) (1.6 mL, 11.6 mmol, 2 equiv) was added and the bomb sealed and heated to 105 °C in an oil bath. The reaction mixture was allowed to stir overnight and upon precipitation of palladium the reaction mixture was cooled and quenched with HCl (0.5 M, 20 mL) and Et_2O (100 mL). The organic layer was washed with H_2O (2x20 mL) and dried with Na_2SO_4 and concentrated. The material was purified via flash column chromatography eluted with EtOAc/Hexanes. The product was isolated as a clear oil. 88% Yield (1.080 g). R_f = 0.5 (33% EtOAc/Hex). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.08 (m, 2H), 7.75 (m, 1H), 7.59 (m, 2H), 2.79 (s, 3H).^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 158.55, 148.88, 130.31, 129.77, 127.44, 125.04, 123.98 (q, *J* = 16), 122.31, 121.29, 119.19 (q, *J* = 5), 25.51. HRMS $C_{11}H_9NF_3$ $(M+H)^+$ calcd. 212.0687, obsvd. 212.0681.

2-(Bromomethyl)-4-(trifluoromethyl)quinoline. 2-Methyl-4-(trifluoromethyl)quinoline (1.000 g, 4.73 mmol, 1 equiv), N-bromosuccinimide (1.690 g, 9.47 mmol, 2, equiv) and azobisisobutyronitrile (77 mg, 0.47 mmol, 0.10 equiv) were all added to a dry rb flask. Then CCl_4 was added and the flask fitted with a condenser and heated to reflux. Careful monitoring by TLC led to quenching after 4 h due to formation of the dibrominated product. The reaction mixture was washed with $NaHCO_3$ solution and then the desired product was separated from the dibrominated product as well as the starting material and the starting material resubjected to the conditions three times to generate adequate material. Colorless Solid. 51% Yield (657mg). R_f = 0.3 (10% EtOAc/Hex). MP = 88-89 °C. ^{1}H-NMR (500 MHz, $CDCl_3$) δ = 8.15 (d, *J* = 8, 1H), 8.11 (d, *J* = 8, 1H), 7.85 (s, 1H), 7.81 (t, *J* = 7, 1H), 7.68 (t, *J* = 7, 1H), 4.72 (s, 3H).^{13}C-NMR (125 MHz,

$CDCl_3$) δ = 156.64, 148.55, 130.92, 130.49, 128.94, 124.51, 124.14, 122.33, 118.60 (*J* = 5.25), 95.00, 33.76. HRMS $C_{11}H_8NF_3Br$ $(M+H)^+$ calcd. 289.9792, obsvd. 289.9789.

2-(Azidomethyl)-4-(trifluoromethyl)quinoline (17). 2-(Bromomethyl)-4-trifluoromethyl) quinoline (530 mg, 1.83 mmol, 1 equiv) was placed in a dry round bottom flask, which was then charged with dry DMF (5.5 ml, 3ml/mmol). Sodium azide (594 mg, 9.14 mmol, 5 equiv) was then added and the reaction flask was fitted with a septum and allowed to stir. After completion via TLC the reaction mixture was quenched with 0.1 M HCl (10 mL) and the aqueous layer extracted with DCM (5 x 15 mL). The organic layers were combined dried over Na_2SO_4 and concentrated. The product was purified using flash column chromatography eluting with EtOAc/Hex. Yellow Oil. 85% Yield (393 mg). R_f = 0.6 (20% EtOAc/Hex). ^{1}H-NMR (500 MHz, $CDCl_3$) δ = 8.11 (d, *J* = 4.5, 1H), 8.06 (d, *J* = 4.5, 1H), 7.74 (t, *J* = 8, 1H), 7.72 (s, 1H), 7.59 (t, *J* = 7, 1H), 4.66 (s, 3H). ^{13}C-NMR {^{1}H} (125 MHz, $CDCl_3$) δ = 155.87, 148.67, 135.53 (q, *J* = 12.8), 130.77, 130.36, 125.56, 124.61, 124.07, 122.42, 116.89 (q, *J* = 2.2), 55.98. HRMS $C_{11}H_8N_4F_3$ $(M+H)^+$ calcd. 253.0701, obsvd. 253.0699.

(4-(Trifluoromethyl)quinolin-2-yl)methanamine (18). In a dry round bottom flask, palladium on carbon (446 mg, .42 mmol (Pd), 0.05 equiv) was suspended in EtOAc (50 mL, 7 mL/mmol). 2-(Azidomethyl)-4-(trifluoromethyl)quinoline was added via a minimal amount of EtOAc and the flask fitted with a balloon of H_2 and purged three times with H_2. After consumption of the starting material by TLC (approx. 15 h), the reaction mixture was filtered through Celite, which was then rinsed with MeOH. The resulting reddish solution was concentrated and taken on without further purification to the general ligand coupling procedure.

(S)-Methyl-2-((4-(trifluoromethyl)quinolin-2-yl)methylcarbamoyl)pyrrolidine-1 carboxylate. See general ligand coupling procedure. Clear Oil. $[\alpha]^{20}_D$ = -52.2° (c = 2.438, $CHCl_3$).

R_f = 0.3 (50% Ace/Hex). Purified in 40% Acetone/Hexane. 21% Yield. NMRs were obtained at higher temperature due to rotamers. ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.07 (t, *J* = 7.8, 2H), 7.74 (t, *J* = 7.6, 1H), 7.60 (m, 2H), 4.75 (s, 2H), 4.41 (m, 1H), 3.66 (s, 3H), 3.55 (m, 2H), 2.30-1.85 (m, 4H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 172.6, 157.0, 148.4, 130.5, 130.0, 128.1, 124.9, 124.1 (d, *J* = 1.2), 122.2 (d, *J* = 4.2), 117.0 (q, *J* = 2.7), 61.1, 52.9, 47.4, 45.1, 29.1, 24.4. Did not observe both carbonyl carbons even at extended relaxation time/high numbers of scans. HRMS $C_{18}H_{19}N_3O_3F_3$ $(M+H)^+$ calcd. 382.1379, obsvd. 382.1375.

(S)-*tert*-Butyl-2-((4-(trifluoromethyl)quinolin-2-yl)methylcarbamoyl)pyrrolidine-1-carboxylate. See general ligand coupling procedure. Clear Oil. $[\alpha]^{20}_D$ = -57.9° (c = 0.676, $CHCl_3$). R_f = 0.5 (50% Ace/Hex). Purified in 20% Acetone/Hexane. 32% Yield. NMRs were obtained at higher temperature due to rotamers. ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.11 (t, *J* = 8.8, 2H), 7.77 (t, *J* = 8, 1H), 7.63 (m, 2H), 4.78 (m, 2H), 4.37 (bs, 1H), 3.52 (m, 2H), 2.38-1.87 (m, 4H), 1.38 (s, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 148.5, 130.5, 130.1, 128.1, 124.9, 124.1 (d, *J* = 2.0), 122.2 (d, *J* = 15), 117.1 (m), 80.6, 61.0, 47.3, 45.2, 28.5, 24.3. Did not observe both carbonyl carbons even at extended relaxation time/high numbers of scans. HRMS $C_{21}H_{24}N_3O_3F_3$ $(M+Na)^+$ calcd. 446.1667, obsvd. 446.1678.

(S)-3-Ethylpentan-3-yl-2-((4-(trifluoromethyl)quinolin-2-yl)methylcarbamoyl)pyrrolidine-1-carboxylate. See general ligand coupling procedure. Yellow Oil. $[\alpha]^{20}_D$ = -53.8° (c = 0.594, $CHCl_3$)R_f = 0.6 (50% Ace/Hex). Purified in 25% Acetone/Hexane. 23% Yield. NMRs were obtained at higher temperature due to rotamers. ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 8.11 (m, 2H), 7.78 (t, *J* = 8, 1H), 7.63 (m, 2H), 4.78 (m, 2H), 4.38 (bs, 1H), 3.53 (m, 2H), 2.29-1.75 (m, 10H), 0.71 (s, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 148.6, 130.6, 130.2, 128.3, 124.9, 124.2, 122.3, 117.3 (m), 115.8, 88.4, 61.0, 47.4, 45.3, 29.9, 27.5, 24.7, 7.9. Did not

observe both carbonyl carbons even at extended relaxation time/high numbers of scans. HRMS $C_{24}H_{30}N_3O_3F_3$ $(M+Na)^+$ calcd. 488.2112, obsvd. 488.2108.

General Propargylation Procedure

To a dry 1.5 dram vial with a Teflon stir bar was added, $CrCl_3(THF)_3$ (18.7 mg, 0.05 mmol, 0.1 equiv), (S)-tert-butyl 2-((4-methoxyquinolin-2-yl)methylcarbamoyl)pyrrolidine-1-carboxylate (21.2 mg, 0.055 mmol, 0.11 equiv), manganese (0) 325 Mesh (54.9 mg, 1 mmol, 2 equiv) and lithium chloride (21 mg, 0.5 mmol, 1 equiv). The vial was then fitted with a cap and Teflon coated septum. The vial was vacuum flushed with argon three times. THF (2.5 mL, 5 mL/mmol) was added followed by TEA (15 *u*L, 0.1 mmol, 0.2 equiv). The reaction mixture was allowed to stir for 3-5 min and then TMSCl (250 *u*L, 2 mmol, 4 equiv) was added dropwise. The reaction mixture was then allowed to stir for 15 min and freshly distilled propargyl chloride was added (110 *u*L, 1.5 mmol, equiv) dropwise. The reaction mixture was then allowed to stir for another 20 min and then acetophenone (59 *u*L, 0.5 mmol, 1 equiv) was added. The reaction mixture was then sealed with Teflon tape and allowed to stir for 72 h. At 72 h, the reaction mixture was quenched by addition of a saturated solution of ammonium chloride in water. The addition of ammonium chloride was slow to allow for effervescence. The mixture was then diluted in diethyl ether and filtered through Celite and the organic layer separated and dried over Na_2SO_4, filtered again and 1M TBAF solution added (0.5 mL). The mixture was stirred for 1-2 h then concentrated and purified via flash column chromatography on neutralized silica gel using EtOAc/Hex or in the case of extremely volatile products Et_2O/Pentane mixtures.

Propargylation Product Characterization

(+)-2-phenylpent-4-yn-2-ol (19a). Known Compound.[5] 75% Yield (61 mg). $[\alpha]^{20}_{D}$ = +29.0° (c = 0.485, $CHCl_3$). Chiral resolution of enantiomers via supercritical fluid chromatography (SFC). ChiraCel AD column, 3 ml/min, 3%IPA, 160 Bar, 40 °C, 6.06 min (minor), 7.06 min (major).

(+)-2-(4-methoxyphenyl)pent-4-yn-2-ol (19b). 74% Yield (71 mg) $[\alpha]^{20}_{D}$ = +23.4° (c = 0.321, $CHCl_3$). Known Compound.[5] Chiral resolution of enantiomers via SFC. ChiraCel AD-H column, 3 ml/min, 3%IPA, 160 Bar, 40 °C, 12.19 min (major), 13.54 min (minor).

(+)-2-(3-(trifluoromethyl)phenyl)pent-4-yn-2-ol (19c). 60% Yield (66 mg). Colorless Oil. $[\alpha]^{20}_{D}$ = +41.3° (c = 0.400, $CHCl_3$). R_f = 0.3 (20% EtOAc/Hex). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 7.75 (s, 1H), 7.64 (d, *J* = 8, 1H), 7.52 (d, *J* = 8, 1H) 7.45 (t, *J* = 8, 1H), 2.70 (m,2H), 2.45 (s, 1H), 2.06 (t, *J* = 2.4, 1H), 1.64 (s, 3H),. ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 147.51, 128.93, 128.43, 124.23 (q, *J* = 3.8), 121.99 (q, *J* = 3.8) 79.92, 73.27, 72.46, 34.80, 29.41. HRMS $C_{12}H_{11}F_3O$ $(M+H)^+$ calcd. 229.0840, obsvd. 229.0842. Chiral resolution of enantiomers via gas chromatography. J&W Scientific 19091G-B213 HP-Chiral-20B cylcodex-B column, 90 °C Isocratic, 26.1 min (minor), 27.5 min (major).

(+)-2-(4-bromophenyl)pent-4-yn-2-ol (19d). 74% Yield (88 mg) $[\alpha]^{20}_{D}$ = +30.5° (c = 0.15, $CHCl_3$). Known Compound.[5] (41) Chiral resolution of enantiomers via SFC. ChiraCel AD-H column, 3 ml/min, 3%IPA, 160 Bar, 40 °C, 15.5 min (major), 16.4 min (minor).

(+)-2-o-tolylpent-4-yn-2-ol (19e). 62% Yield (54 mg). Colorless Oil. $[\alpha]^{20}_{D}$ = +25.3° (c = 0.486, $CHCl_3$). R_f = 0.3 (20% EtOAc/Hex). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 7.44 (m, 1H), 7.16 (m, 3H), 2.84 (ddd, *J* = 39, 8.8, 1.6), 2.57 (s, 3H), 2.37 (s, 1H), 2.05 (t, *J* = 2.4, 1H), 1.69 (s,3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 1.44.28, 128.00, 127.07, 126.19, 80.76, 77.96, 72.05, 37.49, 30.26, 17.78, 17.13. HRMS $C_{12}H_{14}O$ $(M+Na)^+$ calcd. 197.0942, obsvd. 197.0942. Chiral resolution

of enantiomers via SFC. ChiraCel AD-H column, 3 ml/min, 3%IPA, 160 Bar, 40 °C, 5.1 min (major), 6.2 min (minor).

(R)-2-(naphthalen-2-yl)pent-4-yn-2-ol (19f). Known Compound.[5] 72% Yield (76 mg). $[\alpha]^{20}_{D}$ = +34.1° (c = 0.962, $CHCl_3$). Chiral resolution of enantiomers via SFC. ChiraCel AD-H column, 3 ml/min, 7%IPA, 160 Bar, 40 °C, 11.0 min (major), 12.4 min (minor). The absolute configuration of this compounds was determined based on comparison to the previous reports.[5]

(+)-3-phenylhex-5-yn-3-ol (19g). Known Compound.[5] (41) 76% Yield (67 mg) $[\alpha]^{20}_{D}$ = +31.4° (c = 0.265, $CHCl_3$). Chiral resolution of enantiomers via SFC. ChiraCel AD-H column, 3 ml/min, 3%IPA, 160 Bar, 40 °C, 5.4 min (major), 6.1 min (minor).

(+)-2-methyl-3-phenylhex-5-yn-3-ol (19h). 80% Yield (75.3 mg). Colorless Oil. $[\alpha]^{20}_{D}$ = +41.0° (c = 0.124, $CHCl_3$). R_f = 0.3 (20% EtOAc/Hex) R_f = 0.3 (10% EtOAc/Hex) ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 7.43-7.39 (m, 2H), 7.35-7.30 (m, 3H), 2.80 (ddd, *J* = 1.4, 2.8, 19.6, 2H), 2.27 (s, 1H), 2.13 (sep, *J* = 6.8, 1H), 1.92 (t, *J* = 2.8, 1H), 0.83 (dd, *J* = 6.8, 19.6; 6H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 1.44.28, 128.00, 127.07, 126.19, 80.76, 77.96, 72.05, 37.49, 30.26, 17.78, 17.13. HRMS $C_{13}H_{16}O$ $(M+Na)^+$ calcd. 211.1099, obsvd. 211.1105. Chiral resolution of enantiomers via SFC. ChiraCel AD-H column, 3 ml/min, 3%IPA, 160 Bar, 40 °C, 4.7 (major), 5.4 min (minor).

(+)-1-(prop-2-ynyl)-1,2,3,4-tetrahydronaphthalen-1-ol (19i). Known Compound.[5] 61% Yield (57 mg). $[\alpha]^{20}_{D}$ = +38.1° (c = 0.472, $CHCl_3$). Chiral resolution of enantiomers via SFC. ChiraCel AD-H column, 3 ml/min, 7%IPA, 160 Bar, 40 °C, 5.65 min (major), 8.4 min (minor).

(+)-2-(thiophen-2-yl)pent-4-yn-2-ol (19j). Known Compound.[5] 68% Yield (57 mg). $[\alpha]^{20}_{D}$ = +16.1° (c = 0.281, $CHCl_3$). Chiral resolution of enantiomers via SFC. ChiraCel AD-H column, 3 ml/min, 1%IPA, 160 Bar, 40 °C, 11.5 min (major), 12.8 min (minor).

(+)-2-(furan-2-yl)pent-4-yn-2-ol (19k). Extremely Volatile! Colorless Oil. 68% Yield (51 mg). $[\alpha]^{20}_{D}$ = +24.2° (c = 0.366, $CHCl_3$). R_f = 0.2 (20% EtOAc/Hex). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 7.34 (dd, *J* = 1, *J* = 0.6, 1H), 6.30 (dd, *J* = 1.8, 1, 1H), 6.27 (dd, *J* = 1.4, 0.6, 1H), 2.75 (ddd, *J* = 24.2, 8.4, 1.3, 2H), 2.45 (s, 1H), 2.04 (t, *J* = 1.3, 1H), 1.63 (s, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 158.03, 142.05, 110.35, 105.36, 79.99, 71.84, 70.43, 32.45, 26.42. HRMS $C_9H_{10}O_2$ $(M+Na)^+$ calcd. 173.0578, obsvd. 173.0589. Chiral resolution of enantiomers via SFC. ChiraCel AD-H column, 2 ml/min, 2%IPA, 160 Bar, 40 °C, 7.7 min (major), 9.7 min (minor).

(+)-2-cyclohexenylpent-4-yn-2-ol (19l). Known Compound.[5] 72% Yield (59 mg). $[\alpha]^{20}_{D}$ = +3.9° (c = 0.768, $CHCl_3$). Chiral resolution of enantiomers via gas chromatography. J&W Scientific 19091G-B213 HP-Chiral-20B cylcodex-B column, 70 °C Isocratic, 100.7 min (minor), 102.1 min (major).

(+)-4-methyloct-1-yn-4-ol (19m). Colorless Oil. Extremely Volatile! 70% Yield (49 mg). $[\alpha]^{20}_{D}$ = +7.7° (c = 0.060, $CHCl_3$). R_f = 0.2 (10% EtOAc/Hex). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 2.33 (t, *J* = 2.4, 2H), 2.04 (t, *J* = 2.4, 1H), 1.78 (s, 1H), 1.54-1.50 (m, 2H), 1.34-1.28 (m, 4H), 1.23 (s, 3H), 0.89 (t, *J* = 2.4, 3H).^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 81.13, 71.83, 71.35, 41.11, 32.57, 26.47, 26.33, 14.22. HRMS $C_9H_{16}O$ $(M+Na)^+$ calcd. 163.1099, obsvd. 163.1113. Chiral resolution of enantiomers via gas chromatography. J&W Scientific 19091G-B213 HP-Chiral-20B cylcodex-B column, 65 °C Isocratic, 15.6 min (minor), 15.7 min (major).

(+)-4,6-dimethylhept-1-yn-4-ol (19n). Extremely Volatile! Known Compound.[5] 70% Yield (49 mg). $[\alpha]^{20}_{D}$ = +1.1° (c = 0.462, $CHCl_3$). Satistifactory separation of enantiomers of the parent alcohol could not be accomplished so derivation to the triflouroacetate was required for baseline resolution of enantiomers. Derivation was accomplished by taking the purified parent alcohol (49 mg, 1 equiv., 0.35 mmol) in dry ether (2 mL, 6 mL/mmol). Pyridine was added by syringe (90 *u*l, 2 equiv., 0.7 mmol) and cooled to 0 °C in an ice bath. Finally, triflouroacetic

anhydride (85 *u*l, 1.5 equiv. 0.53 mmol) was added dropwise and the reaction allowed to stir at 0 °C for 30 min. The reaction was quenched with water (3 mL) and extracted and analyzed for enantiomeric excess without further purification. Chiral resolution of enantiomers via gas chromatography. J&W Scientific 19091G-B213 HP-Chiral-20B cylcodex-B column, 50 °C Isocratic, 29.1 min (minor), 30.3 min (major).

(+)-2-cyclohexylpent-4-yn-2-ol (19o). Colorless Oil. 62% Yield (52 mg). $[\alpha]^{20}_{D}$ = +8.0° (c = 0.552, $CHCl_3$). R_f = 0.3 (10% EtOAc/Hex). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 2.35 (dd, *J* = 15.2, 2.4; 2H), 2.04 (t, *J* = 2.4, 1H), 1.9-1.6 (m, 6H), 1.52-1.46 (m, 1H), 1.28-0.90 (m, 9H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 81.16, 73.58, 71.54, 48.89, 30.88, 27.90, 27.10, 26.83, 26.76, 26.61, 23.58. HRMS $C_{11}H_{18}O$ $(M+Na)^+$ calcd. 189.1255, obsvd. 189.1265. Chiral resolution of enantiomers via gas chromatography. J&W Scientific 19091G-B213 HP-Chiral-20B cylcodex-B column, 70 °C Isocratic, 105.4 min (minor), 106.2 min (major).

(+)-2,2,3-trimethylhex-5-yn-3-ol (19p). Extremely Volatile! Known Compound.[12] 86% Yield (61 mg). $[\alpha]^{20}_{D}$ = +3.4° (c = 0.581, $CHCl_3$). Chiral resolution of enantiomers via SFC. Chiral resolution of enantiomers via gas chromatography. J&W Scientific 19091G-B213 HP-Chiral-20B cylcodex-B column, 60 °C Isocratic, 45.2 min (minor), 46.8 min (major).

(-)-5-methyl-5-(prop-2-ynyl)dihydrofuran-2(3H)-one (19q). Extremely Volatile! 72% Yield (50 mg). Colorless Oil. $[\alpha]^{20}_{D}$ = -7.2° (c = 0.372, $CHCl_3$). R_f = 0.5 (20% EtOAc/Hex). ^{1}H-NMR (400 MHz, $CDCl_3$) δ = 2.70-2.58 (m, 2H), 2.52 (dd, *J* = 4.8, 2.8, 2H), 2.36-2.28 (m, 1H), 2.06-1.98 (m, 2H), 1.46 (s, 3H). ^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$) δ = 176.38, 84.70, 79.01, 77.55, 71.60, 32.25, 31.38, 29.36, 26.35. HRMS $C_8H_{10}O_2$ $(M+Na)^+$ calcd. 161.0578, obsvd. 161.0578. Chiral resolution of enantiomers via gas chromatography. J&W Scientific 19091G-B213 HP-Chiral-20B cylcodex-B column, 85 °C Isocratic, 25.7 min (major), 28.1 min (minor).

3D Free Energy Relationship Model Development

Step #1: Designing the library and choosing the parameters

In creating free-energy relationships using two or more variables, the first challenge is creating a ligand library in accordance with the principles of experimental design to examine the effects of the chosen variables on the response outcome, which for this study was enantioselectivity. The steric and electronic effects were chosen for this study as a result of positive screening experiments in which a single derivative was synthesized and examined for an effect on enantioselectivity. Choosing model variables should be governed by such screening as well as chemical intuition and synthetic practicality. Similar parameterized values, which could be included into related studies, would be hydrophobicity, polarizability, solvent dipole, quadropole moment, pK_a, hydrogen bond length, molar volume, additional electronic effects and additional steric effects or any other parameterizable variable of interest. The measured response is not limited to enantioselectivity but could include important reaction outcomes such as yield, turnover number, cis-trans selectivity, conversion, or any other desirable, measurable reaction outcome. Choice of parameterization is also crucial. In our example, we chose Hammett values to describe the electronic perturbation due to their similarity to our system as well as their demonstrated robustness. Charton values were chosen over similar steric parameters due to their experimental basis (as opposed to molar volume measurements) and the large number of reported values. Once variables (electronic effect of E, steric effect of S) have been identified, which affect the response variable (enantioselectivity), care should be taken to develop a reliable assay for the response variable.

In considering library synthesis, the primary consideration should be developing a library, which effectively spans the factor space or parameterized range. This is a principle of experimental design which should not be overlooked. We previously reported a study in which

one of the primary conclusions was that experimental design principles used in modeling of systems was better utilized as interpolative instead of extrapolative. In this example, we employed the CEt_3 carbamate substituent as the upper steric bound of our library because it provides the largest Charton value for a simple substituent. Similarly, CH_3 was chosen as the lower steric bound because synthesis of any smaller group (particularly H) is not possible. By choosing these bounds for our ligand library, we effectively span the range of synthetically available, measurable carbamate groups. Similarly, the Hammett electronic parameters which were chosen (OCH_3, H, CF_3) also effectively span the synthetically available factor space of electronic perturbations. With the upper and lower bounds in place, the next consideration was which values to include between the upper and lower bounds. Experimental design dictates that in order to achieve maximum statistical significance, the data points must be evenly distributed. Application of this principle to our parameterized values gives rise to many dilemmas. Because both of the values in this study are only measured as discreet quantities, the choice of other substituents is crucial. Not following this principle can unevenly weight the data set and result in less predictive power. Because our library was 3×3, we only had to consider a single middle substituent for both steric and electronic effects. Our steric parameter space was defined from CH_3 (0.52) to CEt_3 (2.38). The middle point between these two bounds lies at 1.45; however, no synthetically reasonable Charton value lies directly on 1.45. The concession we made was to employ the *t*Bu group, which was synthetically easy to access and in our analysis we considered the small bias this would include in our predictions the midpoint. Again, our choice of a middle substituent (H = 0) biases the data slightly but such bias is unavoidable.

Similar consideration must be given to the library size. We have previously reported success modeling 3×3 libraries. We feel these libraries present a good balance between synthetic effort and statistical significance. However, if different variables are being explored

for which variation is less difficult, larger designs might be beneficial. The number of individual data points collected (library constituents) should be equal to or exceed the number of terms incorporated in the base model.

These are all the preliminary considerations for the experimental design driven approach in this work. The next phase was synthesis and evaluation of the library in the NHK propargylation of acetophenone. As mentioned, in order to develop a more relevant ligand a parallel study was conducted using an aliphatic model substrate, 2-hexanone. The results and analysis were similar such that only the acetophenone model will be discussed in detail.

Step # 2: Modeling the Data

To model the data we chose full 3rd order polynomials as the initial function with which to fit the data but analysis is not limited to polynomial models. In order to maximize the statistical significance (predictive power) of these polynomial functions, we normalized the Charton and Hammett parameter values. The normalization according to Equation 3.4 essentially centers the data set about zero in order to minimize the covariance between terms in the model.

$$Z_P = (P - \mu)/s \quad (3.4)$$

Similarly, having chosen the electronic bounds of our library (OMe = -0.27, CF_3 = 0.54) no Hammett parameter lies close to Z_P is the normalized parameter value, P is the original parameter value and μ is the mean of parameters used in the data set and s is the standard deviation of the parameters used in the data set. Normalization of the both the electronic and steric parameters used in this report yielded Table 3.2.

From this point forward in the discussion, we will use the normalized values as the representative values for the substituents they describe.

Table 3.2. Parameter normalizations.

E	S	*Hammett-σ Value*	*Charton Value*	*Hammett Normalized Values*	*Charton Normalized Values*
CF_3	CH_3	0.54	0.52	1.299	-1.091
CF_3	tBu	0.54	1.24	1.299	-0.178
CF_3	CEt_3	0.54	2.38	1.299	1.269
H	CH_3	0	0.52	-0.260	-1.091
H	tBu	0	1.24	-0.260	-0.178
H	CEt_3	0	2.38	-0.260	1.269
OCH_3	CH_3	-0.27	0.52	-1.039	-1.091
OCH_3	tBu	-0.27	1.24	-1.039	-0.178
OCH_3	CEt_3	-0.27	2.38	-1.039	1.269
				1.299	-1.091

The design matrix is defined as the matrix of parameter coefficients. As an example of how this matrix is assembled we will use the data point given by the ligand E = CF_3, S = CH_3. The translated parameter values for these substituents are CF_3 = 1.299 and CH_3 = -1.091. The response for this value was measured as a 44.6 : 55.4 ratio of minor to major enantiomer. This was converted to $\Delta\Delta G^{\ddagger}$ through the Equation 3.5.

$$\Delta\Delta G^{\ddagger} = -RTln(k_{ent}) \quad (3.5)$$

In Equation 3.5 R is the ideal gas constant given as 0.001986 kcal*$K^{-1}$$mol^{-1}$ and T being 295 for all reported reactions and k_{ent} the ratio of major/minor percent peak areas as measured by SFC chiral resolution of product enantiomers. Thus for our example point E = CF_3, S = CH_3, we calculate $\Delta\Delta G^{\ddagger}$= -0.127 kcal/mol. Using these data we can assemble a row in our design matrix. Using Equation 3.2 as the base model, the design matrix is developed through expansion of our translated parameter values according to the individual terms. For our example point, its corresponding row in the design matrix shown in Table 3.3. Where the values included are the actual values from performing the term operation. z0 is an offset and thus its coefficient is always equal to 1, and for E the parameter coefficient is the value of E (1.299), which produced the corresponding enantioselectivity. The value for E^2 is simply $(1.299)^2$ = 1.687 and the value for ES = 1.299 * -1.091 = 1.687. Using all of the ligands and their replicate runs in this manner constructs the entire design matrix, in Table 3.4. Each row corresponds to a measured response which defines the response matrix Y, given in Table 3.5. The formulation of the design matrix X and the response matrix Y allow us to use the linear algebra definition of linear least squares regression (Equation 3.6) to solve for the matrix of coefficient values C (z0, a, b, c, d, e, f, g, h, i, j).

$$C = (X^T X)^{-1}(X^T Y) \quad (3.6)$$

Table 3.3. Example row of design matrix X.

z0	E	S	E^2	S^2	ES	E^3	S^3	SE^2	ES^2
1	1.299	-1.091	1.687	1.191	-1.417	2.190	-1.300	-1.841	1.547

Table 3.4. Design matrix X.

z0	E	S	E^2	S^2	ES	E^3	S^3	SE^2	ES^2
1	1.299	-1.091	1.687	1.191	-1.417	2.190	-1.300	-1.841	1.547
1	1.299	-0.178	1.687	0.032	-0.231	2.190	-0.006	-0.300	0.041
1	1.299	1.269	1.687	1.611	1.648	2.190	2.044	2.140	2.092
1	-0.260	-1.091	0.067	1.191	0.283	-0.018	-1.300	-0.074	-0.309
1	-0.260	-0.178	0.067	0.032	0.046	-0.018	-0.006	-0.012	-0.008
1	-0.260	1.269	0.067	1.611	-0.330	-0.018	2.044	0.086	-0.418
1	-1.039	-1.091	1.079	1.191	1.134	-1.121	-1.300	-1.178	-1.238
1	-1.039	-0.178	1.079	0.032	0.185	-1.121	-0.006	-0.192	-0.033
1	-1.039	1.269	1.079	1.611	-1.318	-1.121	2.044	1.370	-1.673
1	1.299	-1.091	1.687	1.191	-1.417	2.190	-1.300	-1.841	1.547
1	1.299	-0.178	1.687	0.032	-0.231	2.190	-0.006	-0.300	0.041
1	1.299	1.269	1.687	1.611	1.648	2.190	2.044	2.140	2.092
1	-0.260	-1.091	0.067	1.191	0.283	-0.018	-1.300	-0.074	-0.309
1	-0.260	-0.178	0.067	0.032	0.046	-0.018	-0.006	-0.012	-0.008
1	-0.260	1.269	0.067	1.611	-0.330	-0.018	2.044	0.086	-0.418
1	-1.039	-1.091	1.079	1.191	1.134	-1.121	-1.300	-1.178	-1.238
1	-1.039	-0.178	1.079	0.032	0.185	-1.121	-0.006	-0.192	-0.033
1	-1.039	1.269	1.079	1.611	-1.318	-1.121	2.044	1.370	-1.673

Table 3.5. Response matrix Y.

Average $\Delta\Delta G^{\ddagger}$
-0.1270
0.0141
0.2375
0.0704
1.1034
0.6882
0.2424
1.3413
0.8085
-0.0704
0.2868
0.2868
0.1176
1.1401
0.7692
0.2375
1.4309
0.8884

Two methods are available to iteratively arrive at matrix C. The first and easiest is to enter these matrices into Matlab™ and use the stepwise regression protocol (stepwise(X,Y)). Opening this tool gives the prompt in Figure 3.11. This window allows you to see the estimated coefficient values in the Coeff. column their t-statistics as well as their p-statistic. In the middle of the window it gives the z0 value as Intercept. RMSE is the root mean square estimate; R-squared is the correlation coefficient as well as its adjusted value. F is the Fisher statistic value calculated for the present model and p is the total p-statistic. The bottom dialog box is the model history in terms of RMSE, which allows visualization of how the RMSE changes as you add and omit terms from the top box. In the coefficient box at the top, when the values are indicated in red they are not included into the model and when they are given in blue they are included in the model. The above box shows all of the coefficients not in the model. To conduct a backwards elimination regression each coefficient must be selected as shown in Figure 3.12. With each coefficient selected, clicking the *All-Steps* button iteratively removes terms from the model until it converges into the most statistically significant model as shown in Figure 3.13. Following this protocol we arrive at the model as defined by the coefficients in blue and the intercept indicated in Equation 3.7.

$$\Delta\Delta G^{\ddagger} = 0.955 - 0.568E + 0.307S - 0.454S^2 - 0.097ES + 0.269ES^2 \quad (3.7)$$

Following this iterative protocol can lead to local maxima and care must be taken to ensure that a global maximum is reached. The coefficients reported are different from those in Equation 3.4 due to the normalization of the parameters used. The normalized equation can be decoded through the linear algebra manipulation given by Equation 3.8.

$$B = A^{-1*}C \quad (3.8)$$

where B is the matrix of decoded coefficients C is the matrix of coded coefficients given in Table 3.6.

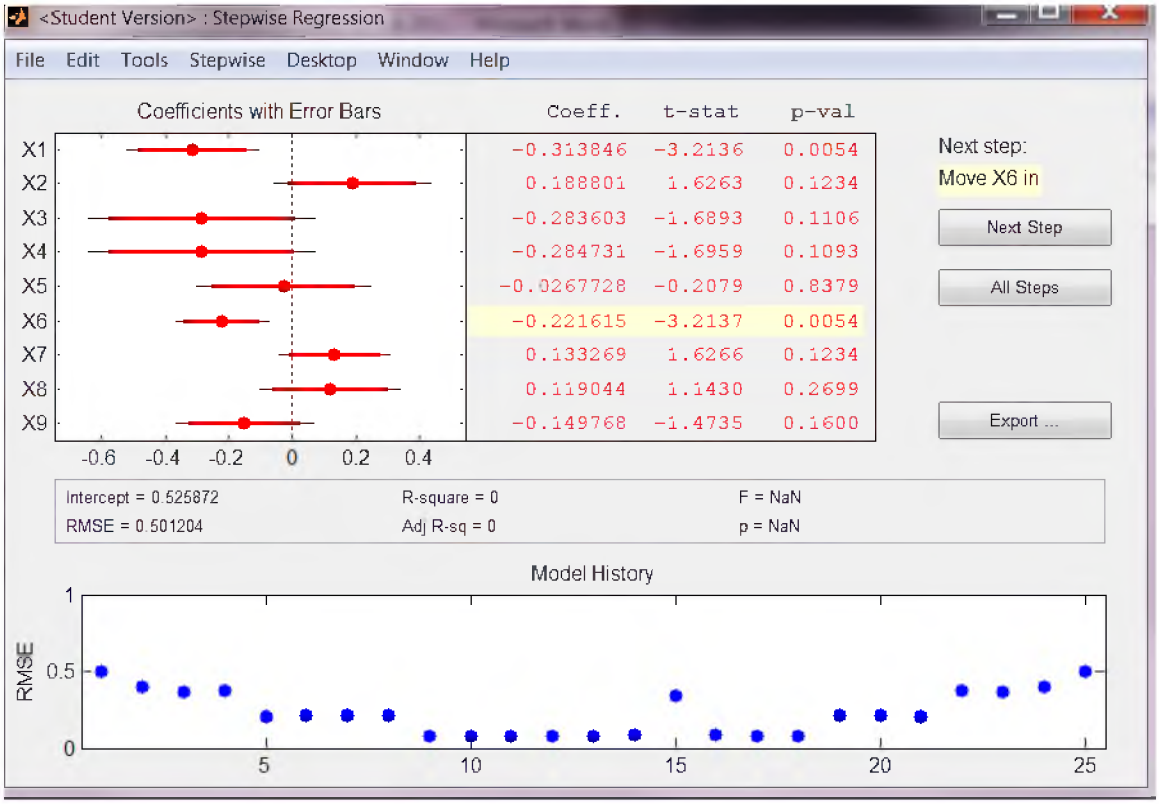


Figure 3.11. Matlab regression solver prompt.

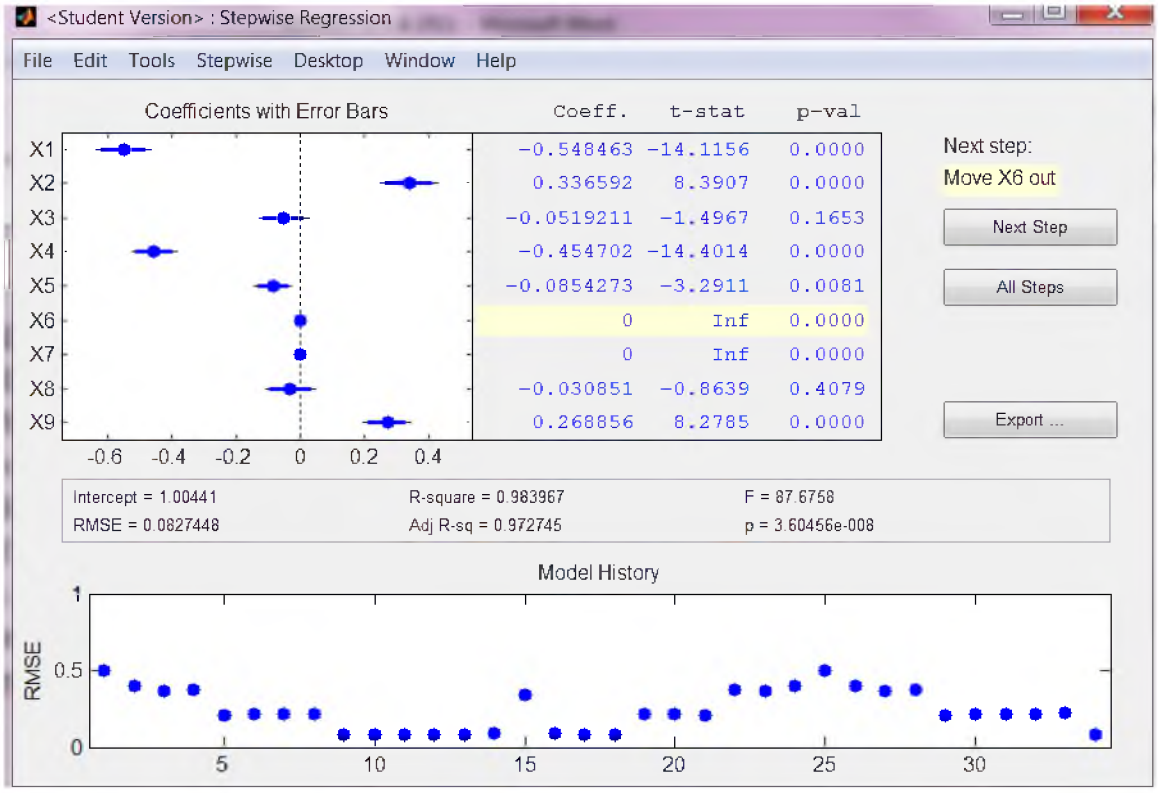


Figure 3.12. Matlab regression solver with all included variables.

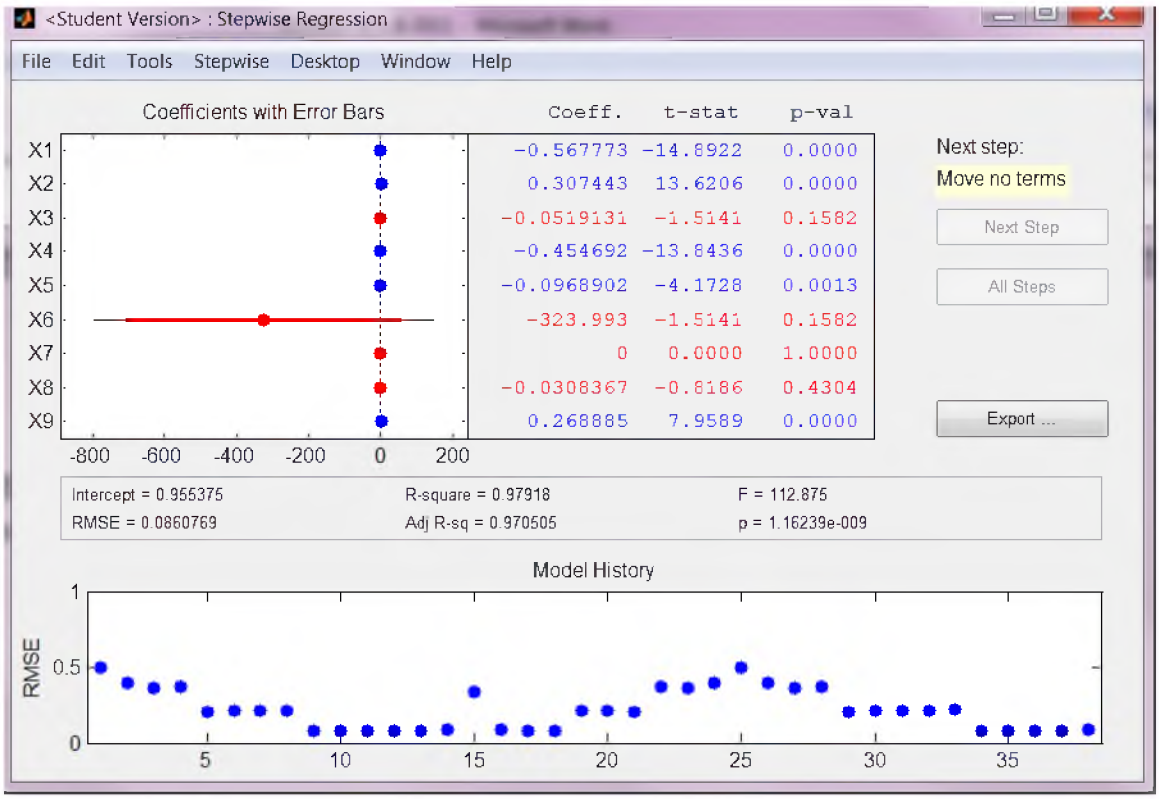


Figure 3.13. The final model as solved by Matlab.

Table 3.6. Matrix of estimated parameter values C.

C
0.95
-0.56
0.37
-0.45
-0.09
0.26

A is the decoding matrix obtained Equation 3.9.

$$A = (X'*X)^{-1} (X' * D) \quad (3.9)$$

where X is the final design matrix given in Table 3.7. D is the decoded design matrix shown in Table 3.8. Manipulations yields matrix B shown in Table 3.9.

LOO Validation of the Model

Validation was performed using the normalized data matrices. Initially, the data was assembled into the final model matrices. Then a single ligand and replicate run was omitted from the training set (Table 3.10, outlined boxes). The remaining data was used to regressed according to Equation 3.5, and new parameter coefficients were generated. The new coefficients were then used to predict the value of the omitted point and the results tabulated.

References

(1) Harper, K. C.; Sigman, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 2179.

(2) Harper, K. C.; Sigman, M. S. *Science* **2011**, *333*, 1875.

(3) Keck, G. E.; Krishnamurthy, D.; Chen, X. *Tetrahedron Lett.* **1994**, *35*, 8323.

(4) Hernandez, E.; Burgos, C. H.; Alicea, E.; Soderquist, J. A. *Org. Lett.* **2006**, *8*, 4089.

(5) Shi, S.-L.; Xu, L.-W.; Oisaki, K.; Kanai, M.; Shibasaki, M. *J. Am. Chem. Soc.* **2010**, *132*, 6638.

(6) Fandrick, K. R.; Fandrick, D. R.; Reeves, J. T.; Gao, J.; Ma, S.; Li, W.; Lee, H.; Grinberg, N.; Lu, B.; Senanayake, C. H. *J. Am. Chem. Soc.* **2011**, *133*, 10332.

(7) Barnett, D. S.; Schaus, S. E. *Org. Lett.* **2011**, *13*, 4020.

(8) Liu, S.; Kim, J. T.; Dong, C.-G.; Kishi, Y. *Org. Lett.* **2009**, *11*, 4520.

(9) Miller, J. J.; Sigman, M. S. *J. Am. Chem. Soc.* **2007**, *129*, 2752.

Table 3.7. Finalized design matrix X.

z0	E	S	S^2	ES	ES^2
1	1.299	-1.091	1.191	-1.417	1.547
1	1.299	-0.178	0.032	-0.231	0.041
1	1.299	1.269	1.611	1.648	2.092
1	-0.260	-1.091	1.191	0.283	-0.309
1	-0.260	-0.178	0.032	0.046	-0.008
1	-0.260	1.269	1.611	-0.330	-0.418
1	-1.039	-1.091	1.191	1.134	-1.238
1	-1.039	-0.178	0.032	0.185	-0.033
1	-1.039	1.269	1.611	-1.318	-1.673
1	1.299	-1.091	1.191	-1.417	1.547
1	1.299	-0.178	0.032	-0.231	0.041
1	1.299	1.269	1.611	1.648	2.092
1	-0.260	-1.091	1.191	0.283	-0.309
1	-0.260	-0.178	0.032	0.046	-0.008
1	-0.260	1.269	1.611	-0.330	-0.418
1	-1.039	-1.091	1.191	1.134	-1.238
1	-1.039	-0.178	0.032	0.185	-0.033
1	-1.039	1.269	1.611	-1.318	-1.673

Table 3.8. Design matrix D based on X with nonnormalized parameter values.

z0	E	S	S^2	ES	ES^2
1	0.540	0.520	0.270	0.281	0.146
1	0.540	1.240	1.538	0.670	0.830
1	0.540	2.380	5.664	1.285	3.059
1	0.000	0.520	0.270	0.000	0.000
1	0.000	1.240	1.538	0.000	0.000
1	0.000	2.380	5.664	0.000	0.000
1	-0.270	0.520	0.270	-0.140	-0.073
1	-0.270	1.240	1.538	-0.335	-0.415
1	-0.270	2.380	5.664	-0.643	-1.529
1	0.540	0.520	0.270	0.281	0.146
1	0.540	1.240	1.538	0.670	0.830
1	0.540	2.380	5.664	1.285	3.059
1	0.000	0.520	0.270	0.000	0.000
1	0.000	1.240	1.538	0.000	0.000
1	0.000	2.380	5.664	0.000	0.000
1	-0.270	0.520	0.270	-0.140	-0.073
1	-0.270	1.240	1.538	-0.335	-0.415
1	-0.270	2.380	5.664	-0.643	-1.529

Table 3.9. Matrix B of uncoded parameter estimates.

B
-1.20
1.22
2.84
-0.85
-3.79
1.25

Table 3.10. Data for the LOO validaiton.

E	S	z0	E	S	S^2	ES	ES^2
CF_3	CH_3	1	1.299	-1.091	1.191	-1.417	1.547
CF_3	tBu	1	1.299	-0.178	0.032	-0.231	0.041
CF_3	CEt_3	1	1.299	1.269	1.611	1.648	2.092
H	CH_3	1	-0.260	-1.091	1.191	0.283	-0.309
H	tBu	1	-0.260	-0.178	0.032	0.046	-0.008
H	CEt_3	1	-0.260	1.269	1.611	-0.330	-0.418
OCH_3	CH_3	1	-1.039	-1.091	1.191	1.134	-1.238
OCH_3	tBu	1	-1.039	-0.178	0.032	0.185	-0.033
OCH_3	CEt_3	1	-1.039	1.269	1.611	-1.318	-1.673
CF_3	CF_3	1	1.299	-1.091	1.191	-1.417	1.547
CF_3	CF_3	1	1.299	-0.178	0.032	-0.231	0.041
CF_3	CF_3	1	1.299	1.269	1.611	1.648	2.092
H	H	1	-0.260	-1.091	1.191	0.283	-0.309
H	H	1	-0.260	-0.178	0.032	0.046	-0.008
H	H	1	-0.260	1.269	1.611	-0.330	-0.418
OCH_3	OCH_3	1	-1.039	-1.091	1.191	1.134	-1.238
OCH_3	OCH_3	1	-1.039	-0.178	0.032	0.185	-0.033
OCH_3	OCH_3	1	-1.039	1.269	1.611	-1.318	-1.673

(10) Lefebvre, O.; Marull, M.; Schlosser, M. *Eur. J. Org. Chem.* **2003**, *2003*, 2115.

(11) Namba, K.; Kishi, Y. *Org. Lett.* **2004**, *6*, 5031.

CHAPTER 4

MULTIDIMENSIONAL STERIC PARAMETERS IN THE ANALYSIS OF ASYMMETRIC CATALYTIC REACTIONS

Introduction

Attempts to quantify steric hindrance in organic chemistry began in the early 20th century. Once it was generally established that steric hindrance could affect reaction rate, these steric effects began to be examined in depth.[1] The greatest amount of attention was given to the inability to reconcile reaction rate with substituent electronic properties for *ortho*-substituted benzoic acids. Taft's work in the area is widely perceived as a breakthrough, as he recognized the need to delineate electronic effects from steric effects.[2] Since his pioneering work, steric effects have been subjected to several different parameterizations including those by Taft himself. The variety of methods through which these steric parameters have been obtained can lead to uncertainty regarding how and when to apply them.

In the area of asymmetric catalysis, we had success applying Charton parameters but, the observed breaks in correlation discussed in Chapters 1 and 2 were still troubling.[3-6] Using Charton's parameters, strong linear correlation had been observed for simple substituents but efforts to extrapolate these correlations led to breaks in the correlation for both acetophenone and benzaldehyde shown in Figure 1.29. Overcoming the breaks in correlation and accurately predicting enantioselectivity a priori was an underlying goal of the oxazoline-proline ligand

library project as discussed in Chapter 2. Our attempts at prediction using multidimensional models were moderately successful, but physical interpretation of the model was obscured through the incorporation of higher order terms. These models were based on the hypothesis that the breaks were due to a change in the mechanism of asymmetric induction. An alternative hypothesis, we considered congruently, is that the application of Charton's parameters was flawed and the breaks in correlation are based on this flawed application. Examination of QSAR literature along with insightful discussions with Professor Marissa Kozlowski (University of Pennsylvania) led us to a set of parameters developed by Verloop and coworkers, the application of which deepened our understanding of how to apply steric parameters to asymmetric catalysis.[1,7-9]

This chapter reviews and compares steric parameters common in quantitative structure activity relationships (QSAR), and how they might be applied in asymmetric catalysis, as the two fields are undeniably similar. Several of the steric-based LFERs heretofore discussed were reanalyzed using more sophisticated Sterimol parameters developed by Verloop and coworkers.

Comparison of Steric Parameters

Taft's parameters, developed in the 1950s, were experimentally determined and have shown robustness in varied applications.[10-12] Taft developed these parameters in his efforts to delineate steric effects from electronic effects in aliphatic ester hydrolysis, by means analogous to those used to derive Hammett's electronic parameters (Figure 4.1). Taft elegantly hypothesized that, under acid-catalyzed conditions, the preservation of charge through the rate-determining step would negate any inductive or resonance electronic contributions from the R substituent. Thus, any variation in the rate of hydrolysis would be proportional to the steric repulsion of the approaching nucleophile. Taft's original experimental results have been

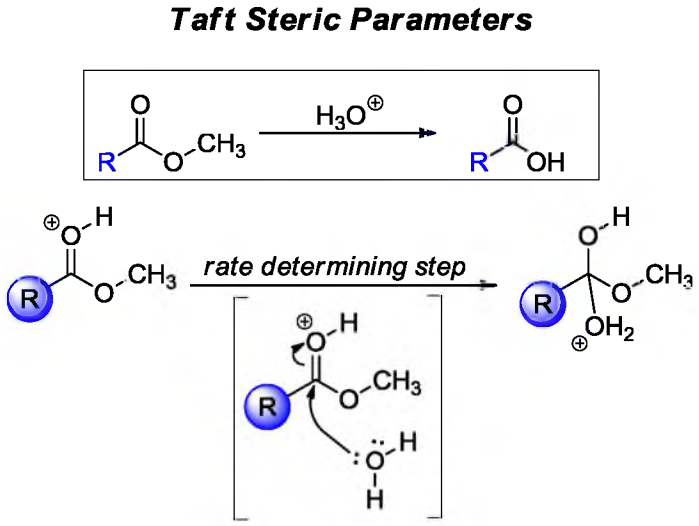


Figure 4.1. The experimental basis of Taft's steric parameter.

manipulated and redefined in many studies, resulting in various sets of Taft-based steric parameters. Most notably, Charton correlated Taft's experimentally measured rates and the calculated minimum van der Waals radii of symmetrical substituents, a relationship that Taft had noted.[13-16] Charton corrected the nonsymmetrical substituents' experimental values to agree with the calculated van der Waals radii, creating a set of computational, but experimentally rationalized, parameters. Hansch validated Charton's parameters by extrapolating Charton's correlation to previously unmeasured substituents, finding agreement between the predicted values and measured values.[17]

Several other steric parameter sets have been developed, both experimentally and computationally, and applied with varying degrees of success in biological and chemical settings (Figure 4.2). Assessing the origin and derivation of some of the most widely known parameters yields insight into how and when they may appropriately be used. Winstein-Holness values (A-values) are probably the most widely recognized set of steric parameters. A-values arise from the conformational study of mono-substituted cyclohexane rings.[18] A-values are based on the observed equilibrium of conformers in mono-substituted cyclohexane rings, where perturbation of this equilibrium can presumably be attributed to 1,3-diaxial steric repulsion. Interference values are another example of an experimentally determined steric parameter, which are based on the heat-induced half-life of racemization in 2,2'-substituted biphenyl systems.[19,20] The steric interaction between the substituent R and opposing aryl ring is presumed to be the key factor responsible for the different energies required for racemization of the atropisomers. Another steric parameter that has had a wide impact on the organometallic community is the Tolman cone angle.[21,22] Tolman and others have measured the projected cone angle of phosphinyl substituents from a hypothetical metal center. Molar refractivity (MR), a steric parameter found

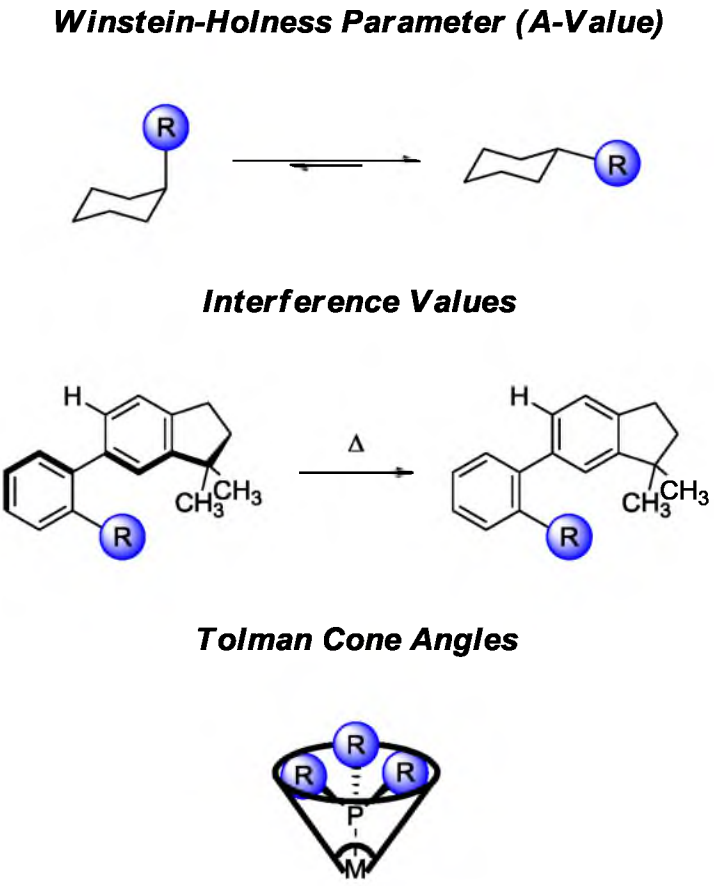


Figure 4.2. Common steric parameters.

in many early QSAR studies, is defined by the Lorentz-Lorenz equation (Equation 4.1) has proven to be an adequate descriptor of total steric volume but disregards molecular shape.[1]

$$MR = [(n^2 - 1)/(n^2 + 2)] * (MW/d) \quad (4.1)$$

Comparison of the normalized (z-score) values derived from the five sets of parameters reveals some innate similarities for the simplest substituents (Figure 4.3). The congruency among interference, A-, Tolman, and Charton values, despite the various means used to define them, suggests these parameters are measuring the same general steric factor. Molar refractivity values do not agree with the other three parameters as closely, but the same general trends are observed.

The broad consistency observed in Figure 4.3 also advocates that selection of the parameter set should be of minimal consequence when examining simple steric effects. The inherent similarity in the parameters suggests that all would be equally robust steric descriptors for simple substituents. Therefore, when selecting a parameter set, the primary consideration would be for which set have the substituents under evaluation been parameterized—not necessarily how the parameters were determined. This consideration is evidenced in the QSAR literature; Taft-based steric parameters are most prevalent, coincident with their extensive parameter library.

Sterimol Parameters

In the 1970s, Verloop and coworkers viewed the similarities between the parameters described above as a fundamental deficiency for describing steric effects. Moreover, application of these simple parameters can fail (vide infra) to provide meaningful steric-based LFERs, most likely due to the multifaceted nature of steric effects. Thus, Verloop developed the Sterimol program, which calculates several dimensional properties for a single substituent, based on

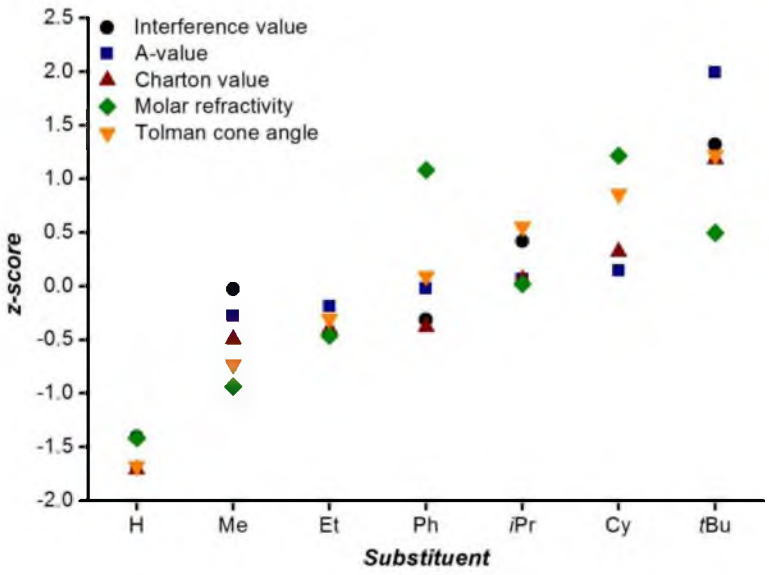


Figure 4.3. Comparison of the normalized values of several parameter sets.

Corey-Pauling-Koltun atomic models.[7-9,23] Rather than group all of the spatial information into a single cumulative value, Verloop created subparameters, each of which describes a different dimensional property of interest (Figure 4.4). Verloop parameters contain three subparameters: two width parameters (B_1 and B_5) and a length parameter (L). The different width subparameters were measured along the substituent's profile when viewed down the axis of the primary bond. The B_1 parameter describes the substituent's minimum profile width orthogonal to the primary bond axis, and the B_5 parameter describes the maximum width also orthogonal to the primary bond axis. The length parameter is the total length of the substituent along the primary bond axis. With B_1 defined as the minimum width perpendicular to the primary bond axis, its value can generally be considered a function of branching at the first carbon center. That is, methyl has a smaller B_1 value than singularly substituted carbons (e.g., ethyl, propyl), and B_1 values get larger with increasing substitution, i.e., disubstituted (e.g., isopropyl) and then quaternary carbons. Verloop saw widespread success in applying these parameters in QSAR studies.

An additional attractive feature of Sterimol parameters is that they are reported in dimensional units (Å), which provides more detailed information about the nature of a steric effect. More specifically, identifying the relationship between enantioselectivity and Sterimol values results in slope coefficients with units of $kcal*mol^{-1}Å^{-1}$. The energy-distance ratio implicates a repulsive steric interaction, which is widely invoked within asymmetric catalysis as a rationale for chirality transfer supporting the potential application of these parameters.

Application of the Charton Parameter to Asymmetric Catalysis

Returning to the previously mentioned breaks in correlation for the NHK reaction shown in Figure 4.5, the central hypothesis that had been put forth was that the breaks in correlation

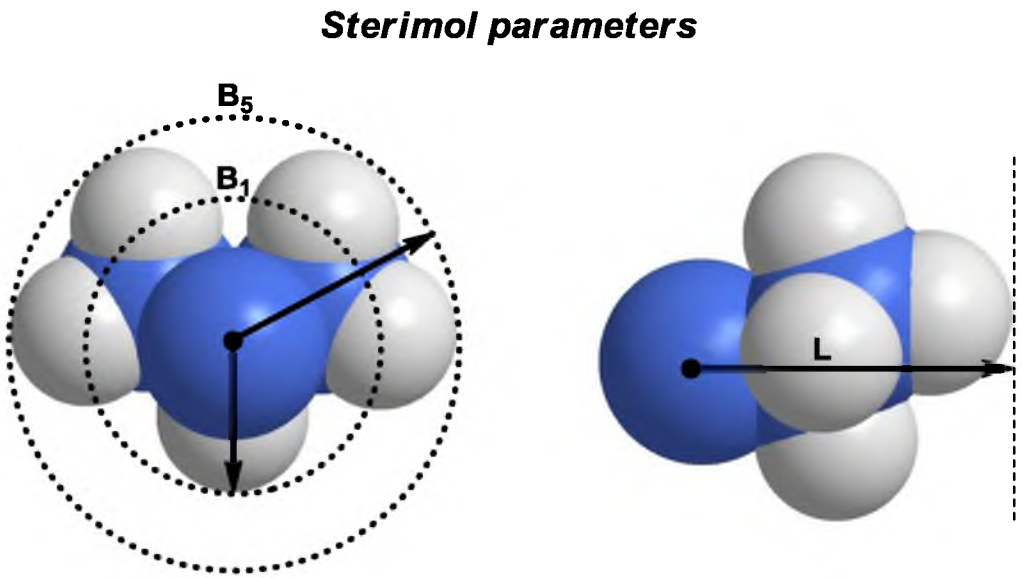


Figure 4.4. Determination of the B_1, B_5 and L Sterimol parameters for an isopropyl group as developed by Verloop and coworkers.

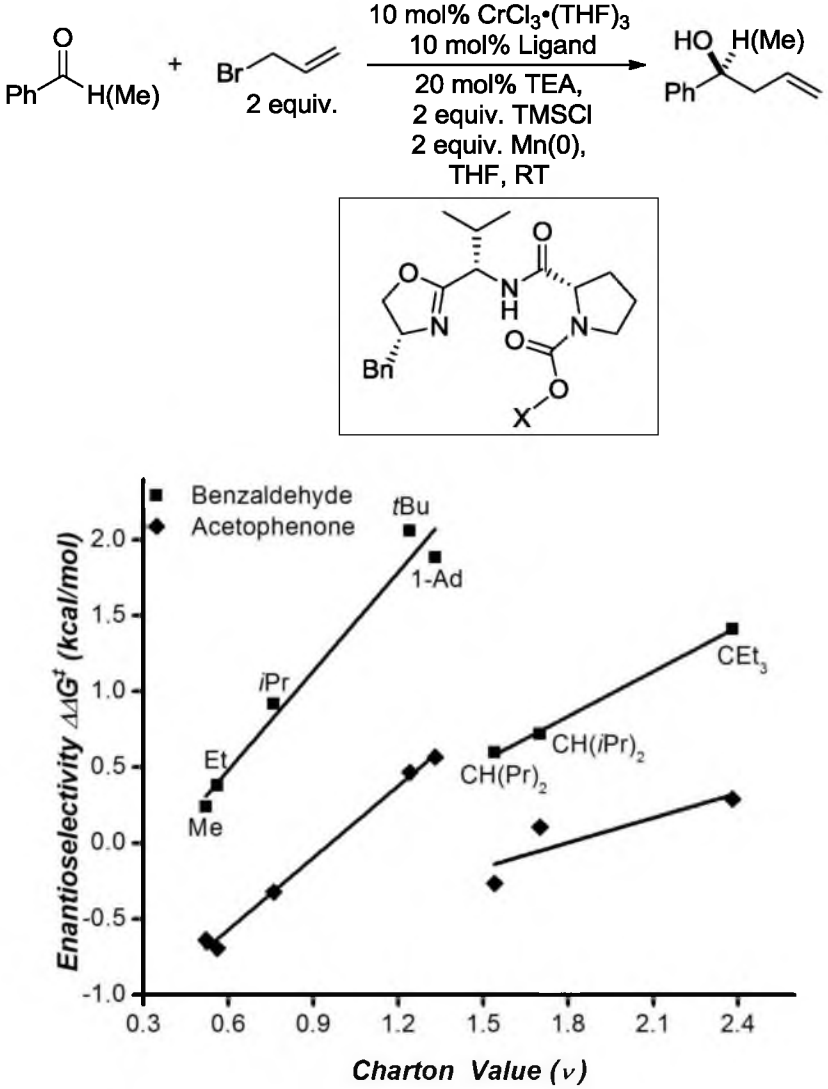


Figure 4.5. The breaks in LFERs for the allylation of acetophenone and benzaldehyde.

represented a change in the mechanism of asymmetric induction.[24,25] Several examples of this interpretation can be found in the literature surrounding Hammett electronic parameters.

However, in that same literature, different parameters have been developed capable of correlating systems where Hammett parameters have failed. Swain-Lupton and Taft-Topsom have reported multidimensional parameterizations of electronic effects adept at correlating electronic effects beyond the scope of Hammett parameters.[17,26,27] As we became familiar with the reasons Verloop articulated for developing Sterimol parameters, we slowly realized that the application of the Charton parameter might be flawed resulting in our perceived breaks in correlation.

To analyze the potential flaw in application, the basis of the Charton parameter must be deconstructed. Charton values are based on the correlation of Taft's the rate of hydrolysis data sets to the minimum Van der Waals radii of substituents. To determine the minimum Van der Waals radius, rotation about the substituent's primary bond is performed, and the minimum radius is the substituent's effective width through this rotation. The result of this treatment is that substituents are generalized as spherical. Therefore, Charton's correlation of minimum van der Waals radii to the relative rate of ester hydrolysis strongly suggests that this bond rotation is kinetically faster than the rate of nucleophilic attack in the hydrolysis reaction. The correlation of van der Waals radii to Taft's experimental data provides compelling evidence for free rotation about the primary bond in acid-catalyzed ester hydrolysis (Figure 4.6). In cases where $R_1 \neq R_2 \neq R_3$, and k_{rot} is much greater than k_{Nu}, then rotation about this bond allows substituents to be described as spheres; in other words, this parameterization implies that the specific bond rotation is significantly faster than the rate-determining step of hydrolysis. When considering an enantioselective reaction under kinetic control, consideration of the Curtin-Hammett principle is critical, wherein all catalyst-substrate conformers, even rotational conformers, do not affect

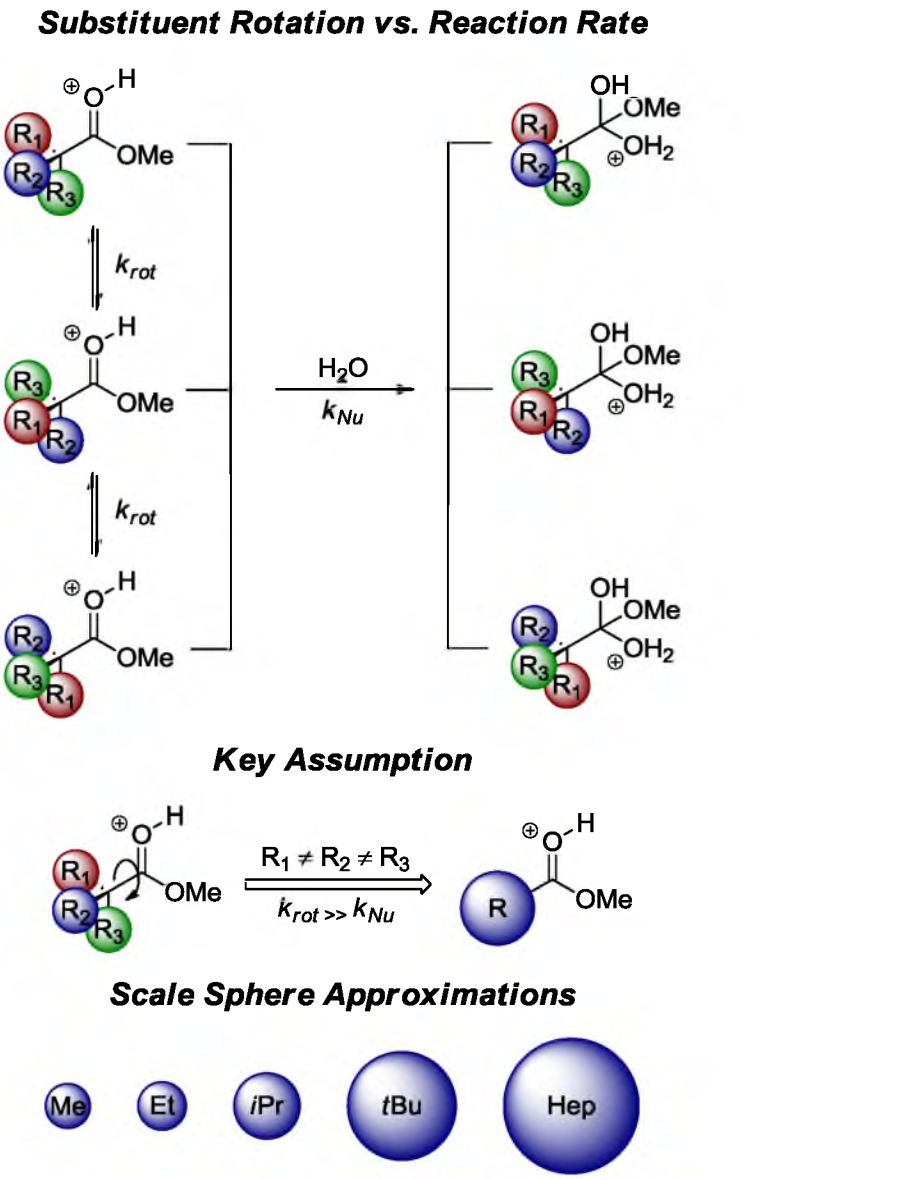


Figure 4.6. The inherent assumptions associated with the Charton parameter and the resulting spherical treatment of substituent which results from this assumption.

enantioselectivity. Only the differential energy associated with the diastereomeric transition states is responsible for enantioselection. Invoking steric effects generally implies that at least one but potentially many possible diastereomeric transitions states are destabilized by repulsive interactions, creating the energy difference resulting in an enantiomeric excess. The steric destabilization of diastereomeric transition states is related to the operating catalyst-substrate conformations approaching the transition state, which is generally simplified to a single low energy conformer.

Juxtaposed to the Curtin-Hammett principle, application of the Charton parameter assumes a net conformer (ball approximation) for a specific transition state, although the difference in energy of these rotational conformers is potentially high. A more probable general scenario is that a single (rotational) conformation provides the lowest energy pathway for nonsymmetrical substituents, and this conformation is not likely to be approximated as a sphere (i.e., $k_{Rot} << k_{Nu}$). In all, Charton's spherical assumption is a limiting premise when the substituent is not symmetrical about the primary bond ($R_1 \neq R_2 \neq R_3$.). For groups with symmetry about the primary bond ($R_1 = R_2 = R_3$), a spherical model based on minimum van der Waals radii more reasonably describes their steric influence in a transition state. However, in our initial studies involving Charton parameters, a linear correlation for nonsymmetrical groups smaller than 1-adamantyl is observed. Why do Charton's parameters seem to fit when smaller substituents are used? A possible explanation could be that the Charton parameters for ethyl and isopropyl are close approximations to the conformationally restricted size, but for larger groups, the disparity between Charton's approximation of the substituent size and the conformationally restricted steric effect is exacerbated.

Sterimol parameters, in contrast, are not based on a mechanistically discrete reaction. The minimum width, as described by the Sterimol parameter B_1, approximates the repulsive

effect of a lowest energy conformer where steric repulsion might occur between a substituent and a substrate. Dependence of enantioselectivity on both minimum and maximum width parameters (B_1 and B_5 respectively) could approximate a less defined transition state and/or indirect steric repulsion and its contribution toward conformational rigidity. A comparison of the normalized values (z-score) of Charton and B_1 parameters highlights the difference between the two sets of parameters, especially for groups larger than cyclohexyl (Cy) (Figure 4.7 and Table 4.1). As with our previous comparison of steric values, the simple, symmetric substituents' values do not differ significantly between the Charton and the B_1 parameters. Examining the larger, nonsymmetrical substituents, the incongruity between the two sets of parameters becomes more apparent, particularly through comparing the order of relative size. The dissimilarities intrigued us so we explored the potential of Sterimol parameters in asymmetric catalysis by reevaluating several of the systems that we previously studied and compare the results.

Analysis of the NHK Allylation Reactions Using Sterimol Parameters

Examination of the data collected for the enantioselective NHK allylation reactions using the Sterimol parameters presented some new challenges in data analysis. Relating the three subparameters to their influence on the enantioselective becomes a four-dimensional problem. Multivariate linear least squares regression techniques similar to those previously outlined were used to develop new models using Matlab™. Parameter coefficients failing the criteria were eliminated from the model until the maximum statistical significance was evolved. The basic model used in the Sterimol reanalysis was a simple first order model shown in Equation 4.2.

$$\Delta\Delta G^{\ddagger} = z0 + aB_1 + bB_5 + cL \quad (4.2)$$

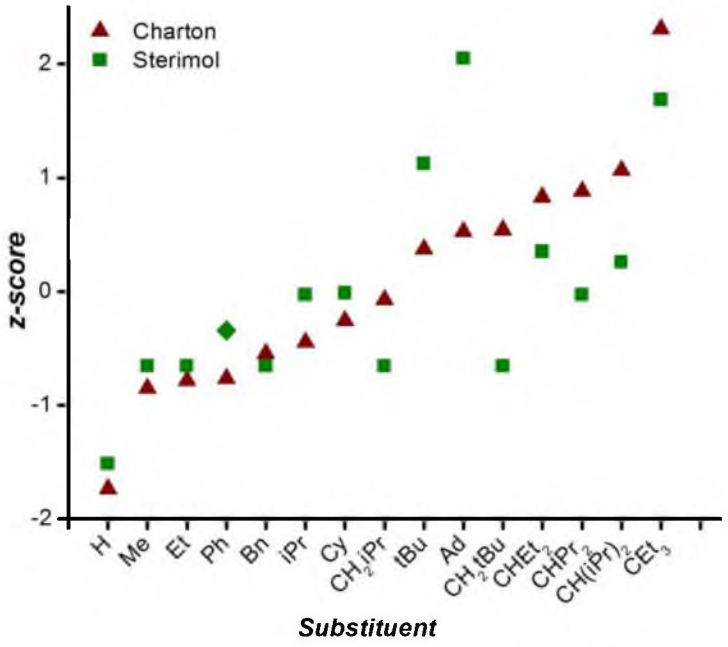


Figure 4.7. Comparison of Charton parameters and Sterimol B_1 values for several substituents.

Table 4.1. Ordering of Charton parameters and Sterimol B_1 values from smallest to largest.

Charton Parameters		***B_1 Sterimol Parameters***	
Group	***Value***	***Group***	***Value***
H	0	H	1
Me	0.52	Me	1.52
Et	0.56	Et	1.52
Ph	0.57	Bn	1.52
Bn	0.7	CH_2*i*Pr	1.52
*i*Pr	0.76	CH_2*t*Bu	1.52
Cy	0.87	Ph	1.71
CH_2*i*Pr	0.98	*i*Pr	1.9
*t*Bu	1.24	$CHPr_2$	1.9
Ad	1.33	Cy	1.91
CH_2*t*Bu	1.34	$CH(iPr)_2$	2.08
$CHEt_2$	1.51	$CHEt_2$	2.13
$CHPr_2$	1.54	*t*Bu	2.6
$CH(iPr)_2$	1.7	CEt_3	2.94
CEt_3	2.38	Ad	3.16

Because of the higher dimensionality of the parameters used, we hypothesized that a simple first-order model could be used to arrive at a meaningful correlation and forgo more complicated higher order terms. Regression of the benzaldehyde NHK allylation data set using Sterimol parameters produced Equation 4.3, which includes only two statistically significant terms: the minimum width parameter (B_1) and the offset (Y intercept) and is plotted in Figure 4.8A.

$$\Delta\Delta G^{\ddagger} = -1.068 + 0.938\ B_1 \quad (4.3)$$

Regression indicated that both the B_5 and L parameters were statistically insignificant. Only the B_1 Sterimol parameter is required to model the data without a break in linearity for larger substituents, unlike the model generated using Charton parameters (Figure 4.8B). Interpretation of the model's dependence on B_1 suggests that steric bulk proximal to the carbamate functionality is crucial for achieving highly enantioselective NHK reactions using this ligand scaffold. This sensitivity towards proximal bulk suggests that the carbamoyl group is closely associated with the catalyst coordination site where facial selection is occurring, possibly through electrostatic attraction. The model also indicates that a group with large proximal steric bulk should generate high enantioselectivity, whereas distal steric bulk (associated with the statistically insignificant B_5 term) would have little effect on the reaction outcome. Accordingly, the NHK reaction employing the 1-adamantyl ligand, which has a large calculated B_1 value, gave the highest observed enantioselectivity.

Comparison of the two LFERs in Figure 4.8 highlights the key difference in the two parameter sets. The treatment of isopropyl-like substituents (CHR_2) varies between the Charton and Sterimol B_1 values. Considering free-rotation about the primary substituent bond, these R-groups inflate the Charton value. The Sterimol calculation of B_1 is similar for the three groups bringing them into correlation with the rest of the observed data.

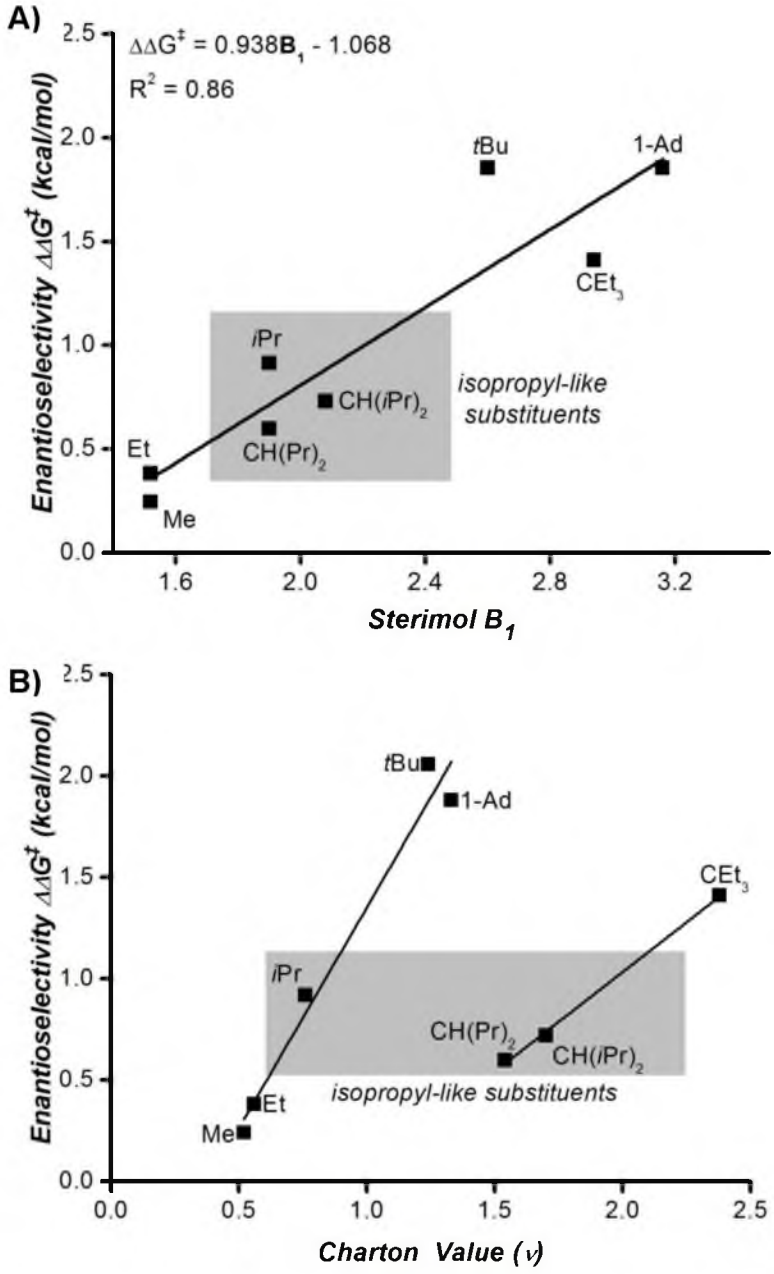


Figure 4.8. A) LFER developed using Sterimol steric parameters for the allylation of benzaldehyde. B) LFER developed using Charton steric parameters for the same reaction.

Reevaluation of the NHK allylation of acetophenone, using Sterimol parameters in the multivariate regression, yielded Equation 4.4.

$$\Delta\Delta G^{\ddagger} = -1.67 + 0.73\ B_1 \quad (4.4)$$

Figure 4.9A depicts how using the Sterimol parameters again provides a linear correlation, in contrast to that observed using Charton parameters (Figure 4.9B). The model with the highest degree of statistical significance uses only the B_1 parameter to correlate enantioselectivity, again indicating close association of the carbamoyl group. Again, the treatment of the isopropyl-like substituents is a key contributor to the observed correlation.

In NHK allylations of both benzaldehyde and acetophenone, the Sterimol parameters provide models with greater correlative power than the Charton-based models. These results also indicate that our previous conclusions based on the Charton analysis were erroneous; a global catalyst change in configuration or activity is no longer consistent with the data. The limitations of the Charton parameters resulted in misrepresentation of the data.

Evaluation of Substrate Steric Effects for the Desymmetrization of Bisphenols

Our lab was recently involved in a collaborative study with the Miller group at Yale University, seeking to elucidate the role of substrate steric effects in the peptide catalyzed desymmeterization of bisphenols, as depicted in Figure 4.10.[28,29] Synthesis of an extensive setof substrate analogs and subsequent evaluation with Charton parameters showed little or no correlation to enantioselectivity (Figure 4.11). This unique example of a steric effect so remote from the bond formation event is perfectly suited to Sterimol analysis, which can attempt to define what key interactions contribute to the observed enantioselectivity. Reevaluating the data using the Sterimol parameters furnished Equation 4.5.

$$\Delta\Delta G^{\ddagger} = -0.418 + 0.929\ B_1 - 0.109\ L \quad (4.5)$$

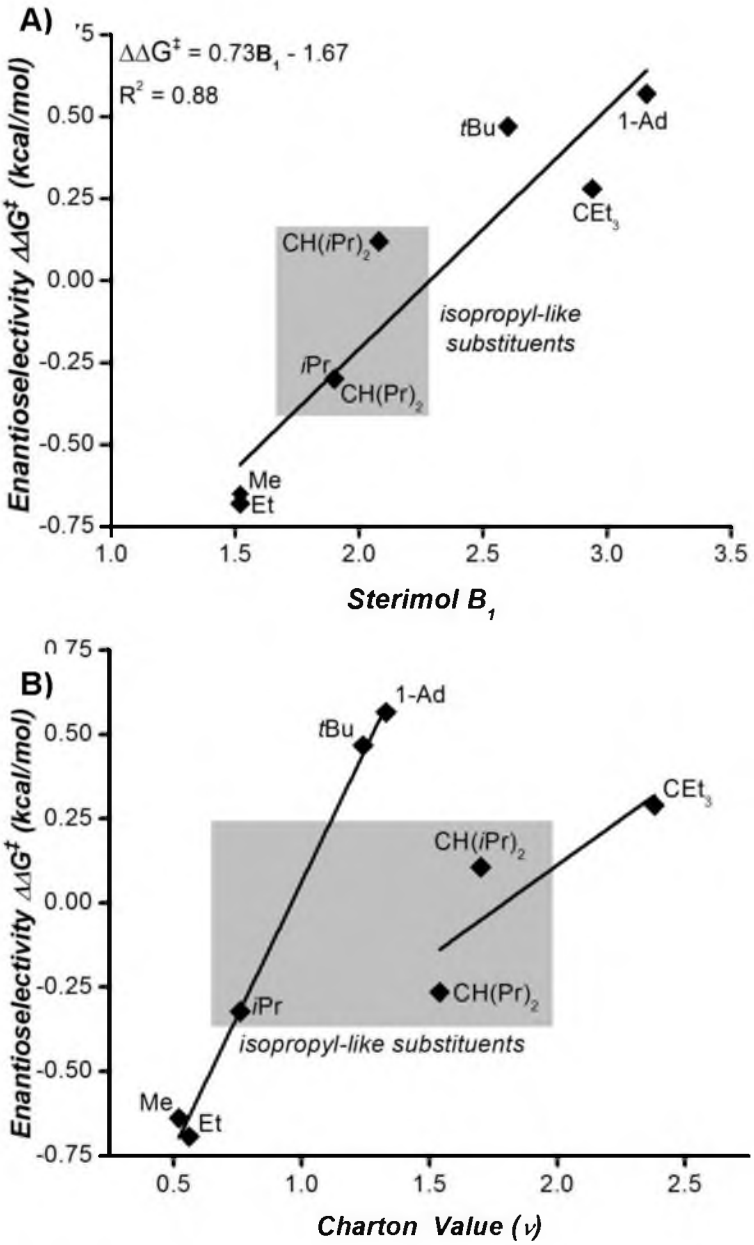


Figure 4.9. A) LFER developed using Sterimol steric parameters for the allylation of acetophenone. B) LFER developed using Charton steric parameters for the same reaction.

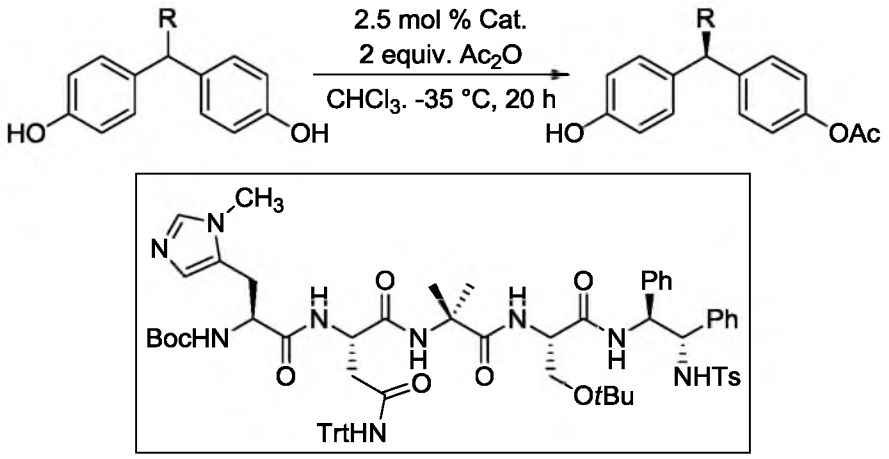


Figure 4.10. The desymmetrization of bisphenols reported by Miller and coworkers.

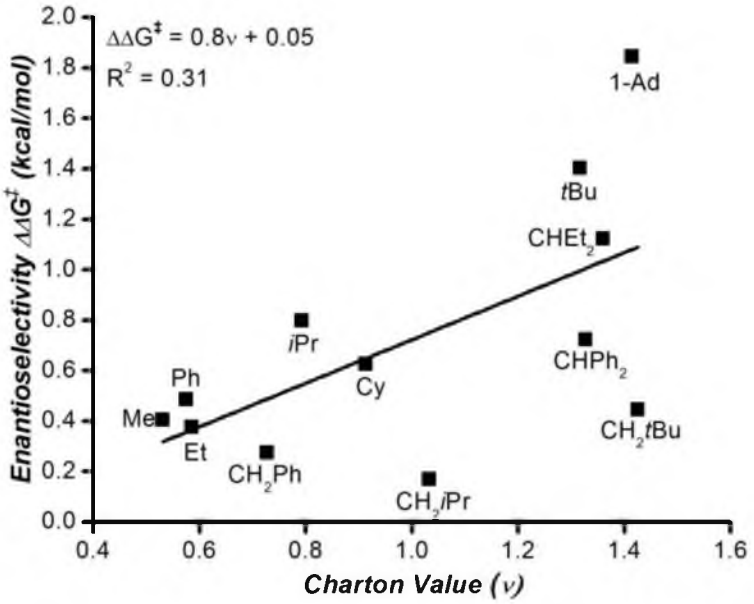


Figure 4.11. Weak correlation observed between the observed enantioselectivity and the Charton value for R in the desymmetrization of bisphenols.

For this reaction, the model dictates the B_5 term as statistically insignificant, and inclusion of B_1 and L terms results in a planar surface relating a substituent's minimum width and length to enantioselectivity (Figure 4.12A). The surface demonstrates that the optimal substrate would possess an R group with a large minimum width and a short length.

The use of Sterimol parameters creates a strong correlative model (Figure 4.12B), further demonstrating the robustness of Sterimol-based models in asymmetric catalytic applications. The model has a high degree of dependence on the B_1 parameter, demonstrating the need for proximal steric bulk. A small but significant negative effect on enantioselectivity is incurred for substituents with larger length parameters. Of the possible mechanisms through which this asymmetric induction may occur, these results seem to implicate the previously hypothesized propeller-like twist of the aryl rings by interaction with a proximally large substituent (Figure 4.13).

All of the available data was employed to derive Equation 4.5, so external validation was unavailable. To validate the model, LOO validation was performed as explained in Chapter 3. The results of this validation are given in Figure 4.14. The high value for Q^2 indicates the model is robust despite the lack of design considerations in the source data.

The strength of the Sterimol parameters in correlating variation in the data suggests that a preferred conformer of the R group in relation to the chiral catalyst exists (i.e., it cannot be treated as dynamic). However, the small negative dependence of the enantioselectivity on the substituent length also indicates that longer R groups will weaken the substrate-catalyst interaction leading to decreased levels of enantioselectivity. This suggests that the substrate is bound such that the R group is arranged toward the catalyst and not away from it. This considerable simplification of the number of possible transition states is not necessarily implied

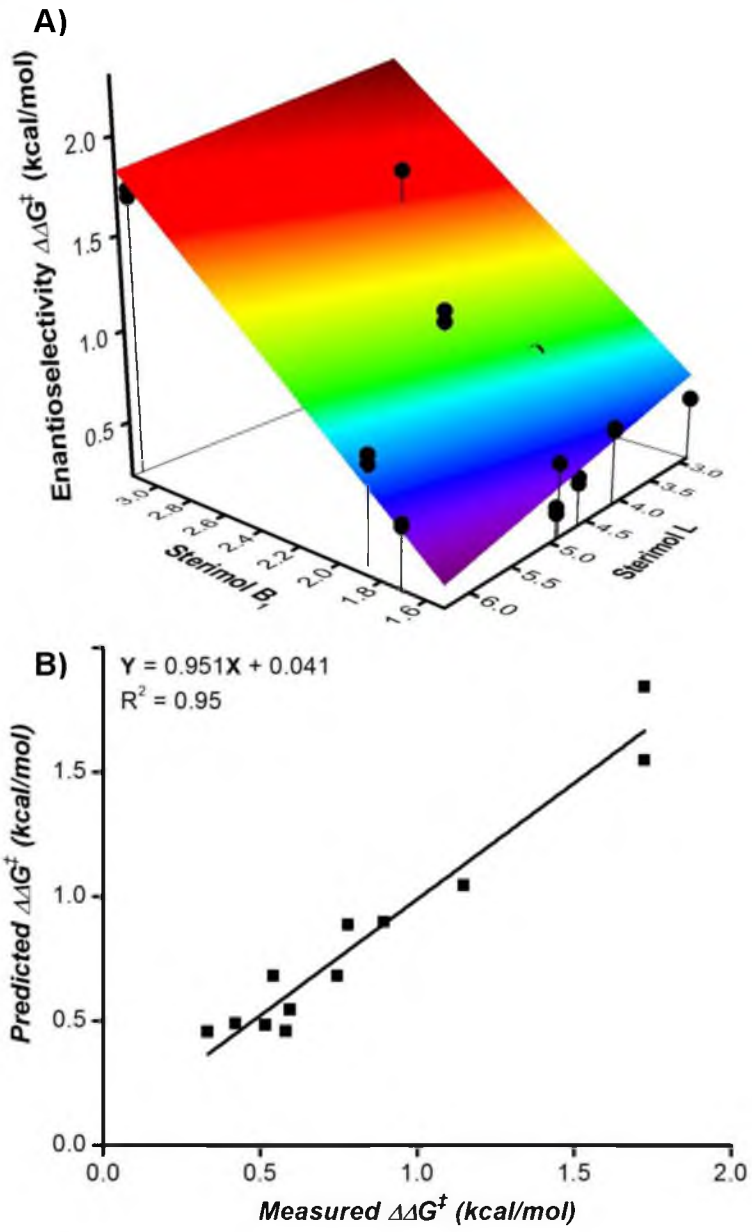


Figure 4.12. A) Surface described by Equation 5. B) Correlation between the values predicted by Equation 5 and the observed enantioselectivities.

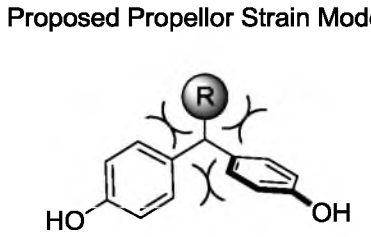


Figure 4.13. Proposed model of propeller-like strain and is effect on enantioselectivity.

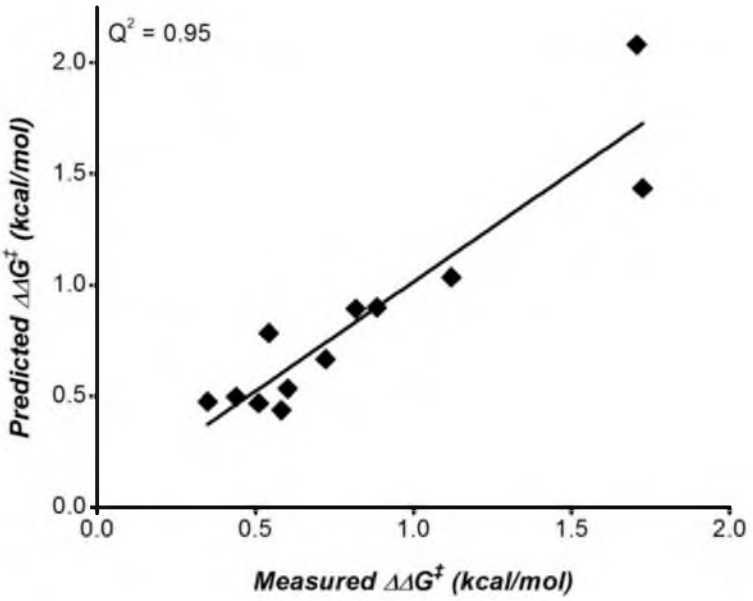


Figure 4.14. Results of LOO validation performed on Equation 5.

solely from the suggested propeller mechanism. The relatively small value of the coefficient for the L parameter advocates that either the R group is some distance away from the catalyst (incurring large repulsive penalties only with extended substituents), or the R group is not oriented directly into the catalyst but angled such that the steric effect is ancillary.

Unlike other steric parameters, Sterimol parameter values are not relative to a hydrogen substituent. Thus, extending the analysis to consider the Y-intercept value can be instructive. A nonzero Y-intercept value indicates that, in the absence of a substituent ($B_1 = 0$, L = 0), the reaction would still proceed with modest enantioselectivity (-0.42 kcal/mol) and opposite facial selectivity; indicating an energetic baseline inherent in the catalyst-substrate interaction. The change in facial selectivity that would arise when a group smaller than hydrogen is imaginarily incorporated at the R position is reasoned to be due to the other methine hydrogen becoming the dominant steric contributor towards the suggested propeller-twist mechanism. The difference in size between the two substituents at the pro–stereogenic center appears to make a significant contribution to the observed enantioselectivity, again suggesting the propeller strain is a contributing factor.

Another potentially interesting tool for delineating catalyst and substrate effects on enantioselectivity is the "H"-intercept, or the value determined by the model when R = H, the smallest possible substituent. For the desymmetrization of bisphenol, this analysis is particularly unique. Inserting the Sterimol values for a hydrogen atom into Equation 4.5 ($B_1 = 1$, L = 2.06), leads to a difference in Gibb's free energy of 0.29 kcal/mol. This energy difference represents an imaginary enantioselectivity because the product of the reaction would be achiral. This computed energy difference indicates a fundamental bias enacted on the substrate by the catalyst. In the transition state, the catalyst-substrate interaction would render the methylene hydrogens diastereotopic, and H-intercept analysis reflects the difference in energies between

the diastereotopic hydrogens in the catalyst-substrate complex. This energy difference is amplified through self-interaction and possibly greater substrate-catalyst repulsion when larger substituents are included. In other analyses, H-intercept evaluation of the substrate could provide insight into baseline energetic bias introduced onto a substrate by a catalyst.

H-intercept analysis is useful when applied to substrate-catalyst interactions, as in the above example, but its extension to the NHK allylation systems previously described becomes complicated by several factors. In the above NHK reactions, the carbamoyl group does not contain a stereocenter; it is simply an extension of the chiral-proline, which could itself induce some enantioselectivity on the reaction and complicate interpretation of the Y- or H-intercepts. Similarly, the NHK catalysts possess several stereocenters, thus the observed steric effect of the carbamate is also related to secondary conformations, as we have previously shown.

Sterimol Analysis of the NHK Propargylation Reaction

Examining the NHK propargylation of ketones discussed in the previous chapter revealed some interesting correlations. We had previously demonstrated a steric-electronic cross relationship via crossterms which gave unique insight into the interplay of these two reaction elements that are traditionally thought of as independent.[6]

Reevaluation of the QuinPro ligand library was performed using Sterimol parameters in place of Charton parameters (Figure 4.15). The base model included all the Sterimol parameters, Hammett values, and crossterms relating each Sterimol subparameter to the Hammett parameter shown in Equation 4.6.

$$\Delta\Delta G^{\ddagger} = z0 + aB_1 + bB_5 + cL + dE + fEB_1 + gEB_5 + hEL \quad (4.6)$$

A backward-elimination stepwise regression using the data and parameter values shown in Table 4.2 was then performed on the system to generate Equation 4.7.

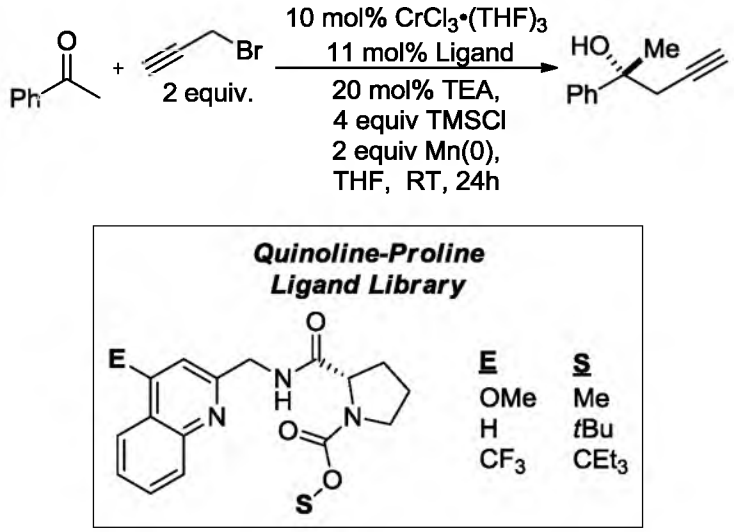


Figure 4.15. The propargylation of acetophenone using the QuinPro library.

Table 4.2. Raw data for the propargylation of acetophenone.

E	*S*	*Hammett-σ Value*	*Charton Value*	*Average er*	*Average ΔΔG‡*
CF_3	CH_3	0.54	0.52	46:54	-0.10
CF_3	*t*Bu	0.54	1.24	56:44	0.15
CF_3	CEt_3	0.54	2.38	61:39	0.26
H	CH_3	0	0.52	54:46	0.09
H	*t*Bu	0	1.24	87:13	1.12
H	CEt_3	0	2.38	78:22	0.73
OCH_3	CH_3	-0.27	0.52	60:40	0.24
OCH_3	*t*Bu	-0.27	1.24	91:9	1.39
OCH_3	CEt_3	-0.27	2.38	81:19	0.85

$$\Delta\Delta G^{\ddagger} = -0.696 + 1.83\ B_1 - 0.962\ B_5 - 2.705\ EB_1 + 1.736\ EB_5 \quad (4.7)$$

The dimensionality of Equation 4.7 prevents visualization of the model but examination of the slope value approaching unity in the predicted versus measured plot (Figure 4.16) shows the strong correlative power of the Sterimol-based model.

In order to determine the predictive power of this model, a cross validation technique was undertaken. All available data were used as a training set to determine Equation 4.7, and in turn, Equation 4.7 provides excellent correlation to the measured responses. To validate this model, another LOO cross-validation was performed as described in Chapter 3. The results of this LOO validation are shown in Figure 4.17. The high Q^2 correlation shown in Figure 4.17 lends credibility to the predictive power of Equation 4.7.

The most interesting feature of Equation 4.7 is that it lacks a solely electronic term and only gives significance to the electronic parameter when paired with a steric component; the electronic contribution to enantioselectivity is only effectual in tandem with the steric effect of the carbamate. This observation can be rationalized by the fact that the electronic component alone does not contain any chiral elements and cannot affect an enantioselective result by itself; only in tandem with the steric effect is enantioselectivity achieved. A possible explanation of this relationship as defined by the Sterimol parameters is that decreasing the Lewis acidity of the chromium (Cr) center is instrumental in decreasing the bonding distance of the carbonyl to the Cr center, thereby amplifying the steric influence of the S group. The model shows a severe penalty for combining substituents with large B_5 values with electron-donating groups. This relationship suggests that larger substituents might congest the reaction at the site responsible for high enantioselectivity, or extended steric repulsion might decrease the proximity of chiral information to the site of enantioselection. In our previous studies of this reaction, we

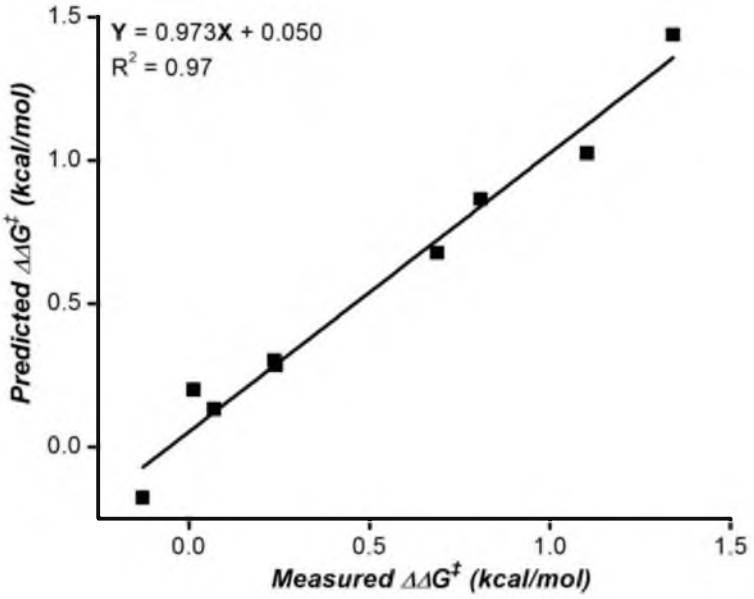


Figure 4.16. Correlation between the values predicted by Equation 7 and the observed enantioselectivities.

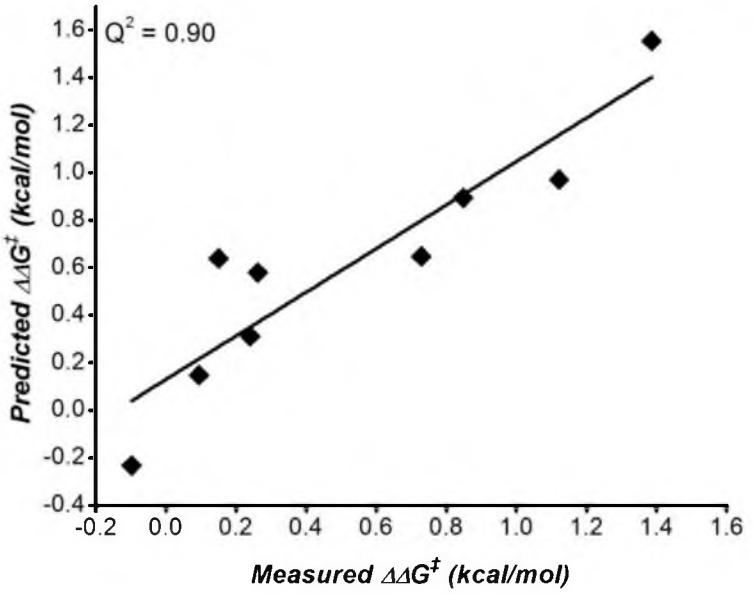


Figure 4.17. Results of the LOO validation on Equation 7.

concluded that the presence of the crossterms within the model gave evidence that steric and electronic effects were acting synergistically to affect enantioselectivity. Use of the more refined Sterimol parameters gives more validity to this relationship and affirms that the steric effect is amplified by the electronic effect and not vice versa.

Deconstructing Equation 4.7 into its individual elements, a strong dependence on the B_1 term is observed, as indicated by its large coefficient. The inclusion of a negative B_5 term indicates that substituents containing a large maximum radius would erode enantioselectivity. The large negative coefficient relating B_1 to E suggests that the positive effects of proximal steric bulk on enantioselectivity are amplified by electron-donating groups (negative Hammett values). The smaller but positive coefficient relating B_5 and E indicates that the opposite effect is observed for large substituents. These two crossterms suggest that the electronic nature of the catalyst influences the transition state, perhaps through a less Lewis-acidic metal, allowing for a more product-like transition state, as alluded to in Jacobsen epoxidation catalysis and alkene difunctionalization reactions developed in our labs.[30,31] This ultimately increases the proximal steric relevance of the S substituent.

The propargylation model was subjected to an optimizing algorithm to predict the key structural elements required for high enantioselectivity. The optimizing routine was bounded by the highest and lowest values contained in the data set for the terms included in the model (B_1, B_5, and E). These bounds are necessitated by synthetic constraints, as well as the inherent problems associated with multidimensional extrapolation. Employing this routine systematically across the ligand space yielded identification of the same optimal ligand structure, indicating that the true global maximum was most likely found. The maximum indicated that the optimal ligand would contain the largest possible B_1 value (2.94), the smallest possible B_5 value (2.04), and the most negative possible Hammett value (-0.27). The required Hammett value would be a

strongly electron-donating group, which agrees with our previous results. The optimized values for B_1 and B_5 are conflicting in terms of synthetically feasible substituents. However, the *t*Bu group presents a reasonable trade-off between a large B_1 and a small B_5 value. Not surprisingly, the optimization of the Sterimol models yields the same optimal ligand as the previous study, with E_{OMe} and S_{tBu}.

Conclusion

Sterimol parameters, as defined by Verloop, present a conceptually different approach to steric parameterization. A-, interference, and Taft-based Charton values are experimentally verified parameterizations but, as such, they are limited by the experimental conditions from which they were obtained. Verloop parameters are computationally-derived and may be a more suitable parameter for describing interactions in asymmetric catalysis due to their increased ability to describe nonsymmetrical steric effects. Implementation of these parameters demonstrated general robustness in comparison to Charton parameters despite the lack of experimental design considerations in developing the models. Future use of experimental design principles congruent with Sterimol parameters could yield more insightful models.

The comparison of models developed with Charton parameters to those developed with Sterimol parameters demonstrates the strength of Sterimol parameters and highlights the detrimental assumptions employed in using Charton parameters. The information-rich models developed with Sterimol parameters for NHK allylations also provide evidence for the specific steric elements influencing enantioselectivity. Analysis of steric effects using Sterimol parameters in the desymmetrization of bisphenol shed insight to mechanism of asymmetric induction and key substrate-catalyst interactions. Similarly, in the reanalysis of data gathered in

the propargylation of acetophenone, the Sterimol derived model delineated which steric effect(s) were amplified and nullified through steric-electronic synergy. Outlined herein were several examples where Sterimol models were insightful and beneficial; however, in our studies, we identified limitations to these parameters and their application, particularly in relation to our previously published steric-steric 3D relationships. Even so, Sterimol parameters, although imperfect, constitute a significant improvement in terms of overall utility and their ability to provide mechanistic insight into reactions. Information of this nature can be applied to developing and verifying stereochemical models, which is the crux of asymmetric catalyst development, and implementation of this information is our continuing goal.

Experimental

Model Development for the Allylation of Benzaldehyde

The benzaldehyde Sterimol model was derived using Matlab™ software. The raw data was compiled into numeric matrices X and Y according to ligand substituent parameters and measured response, respectively, as shown in Table 4.3. Matrices X and Y were input into the Matlab™ software. Next, the stepwise solver was initiated using the command "stepwise(X,Y)". This command produced the prompt window shown in Figure 4.18. Figure 4.18 shows the coefficients for each term (X1, X2, and X3 correspond to B_1, B5, and L, respectively) in the upper right box and also indicates whether each term is (blue) or is not (red) included in the current model. Figure 4.18 also shows that no terms are included into the model. The center box gives the intercept value, as well as selected relevant statistics, for the model given by the blue coefficients in the upper right box. With no terms included in the model, these statistics are meaningless in Figure 4.18. The bottom box is a history of the values of the root-mean-square error (RMSE) of each model calculated. By clicking on each term in the upper right hand

Table 4.3. Matrix formatting of the benzaldehyde data.

	Matrix X			Matrix Y
Substituent	B_1	B_5	L	Average $\Delta\Delta G^{\ddagger}$
Me	1.52	2.04	2.87	0.245
Et	1.52	3.17	4.11	0.381
*i*Pr	1.9	3.17	4.11	0.913
*t*Bu	2.6	3.17	4.11	1.854
$CH(Pr)_2$	1.9	4.54	6.17	0.597
CEt_3	2.94	4.18	4.92	1.410
$CH(iPr)_2$	2.08	4.19	4.12	0.730
1-Ad	3.16	3.49	6.17	1.854

Figure 4.18. The stepwise regression prompt.

box with red and blue terms, each term is incorporated into the model to give the display shown in Figure 4.19.

In Figure 4.19, note that all of the terms are shown in blue, which means all terms are included in the present model. The "Model History" now has four points for the separate inclusion of each of the three terms. The coefficients are estimated for the B_1, B_5 and L terms in the upper right hand box. These values are subjected to t- and p-tests to measure their significance in the model at hand. These statistics show that the L term is insignificant and should be eliminated from the model (as prompted by the "Next Step" indicator). Removal of this term leads to the model shown in Figure 4.19. Figure 4.20 shows the model that excludes the L term. Again, analysis of the statistical measures for this model prompts us to eliminate the B_5 term. Figure 4.21 shows the model that contains only the B_1 term and an intercept value shown in Equation 4.3. Using this model, each of the synthetic ligands' enantioselectivities were predicted (Table 4.4).

Cross Validation for the NHK Allylation of Benzaldehyde Model

To validate the model given in Equation 4.3, a simple cross validation was performed. All data collected for the reaction were used in the training set to develop Equation 4.3. To cross validate this model as a measure of its true predictive power, a Leave-One-Out (LOO) validation was performed. These models were constrained to fit a model using only the terms of the base model of Equation 4.3 (z0, B_1) using regression techniques outlined above. These models were then used to independently predict the omitted data point as given in Table 4.5.

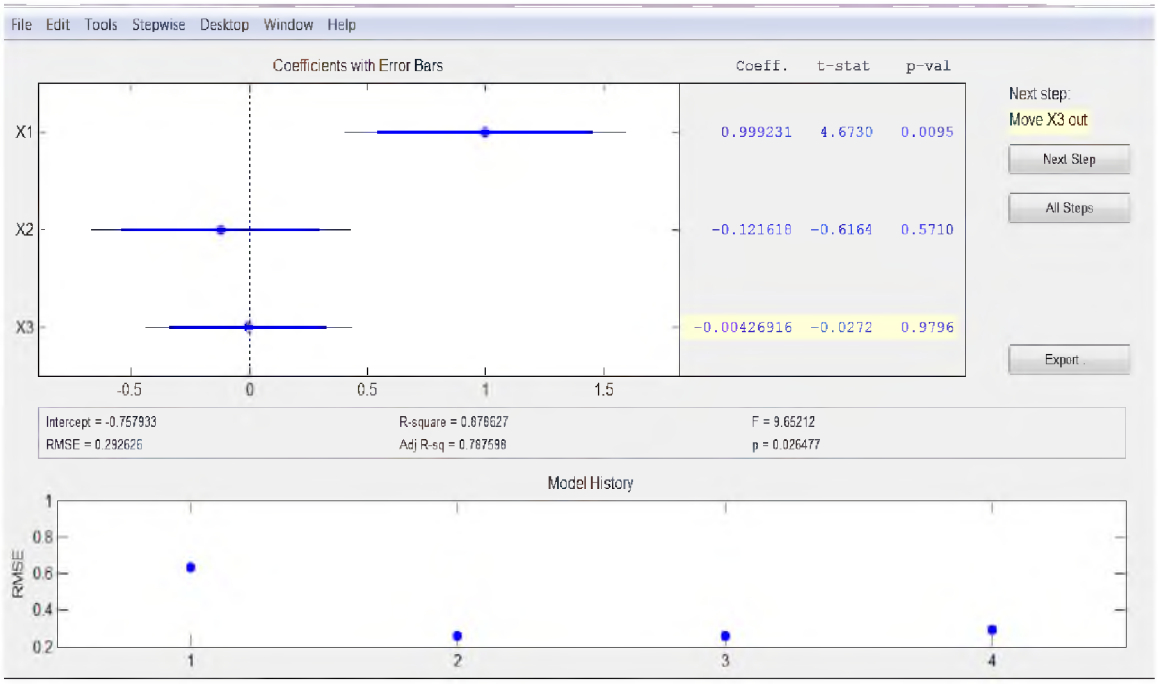


Figure 4.19. Initial iteration in the benzaldehyde regression.

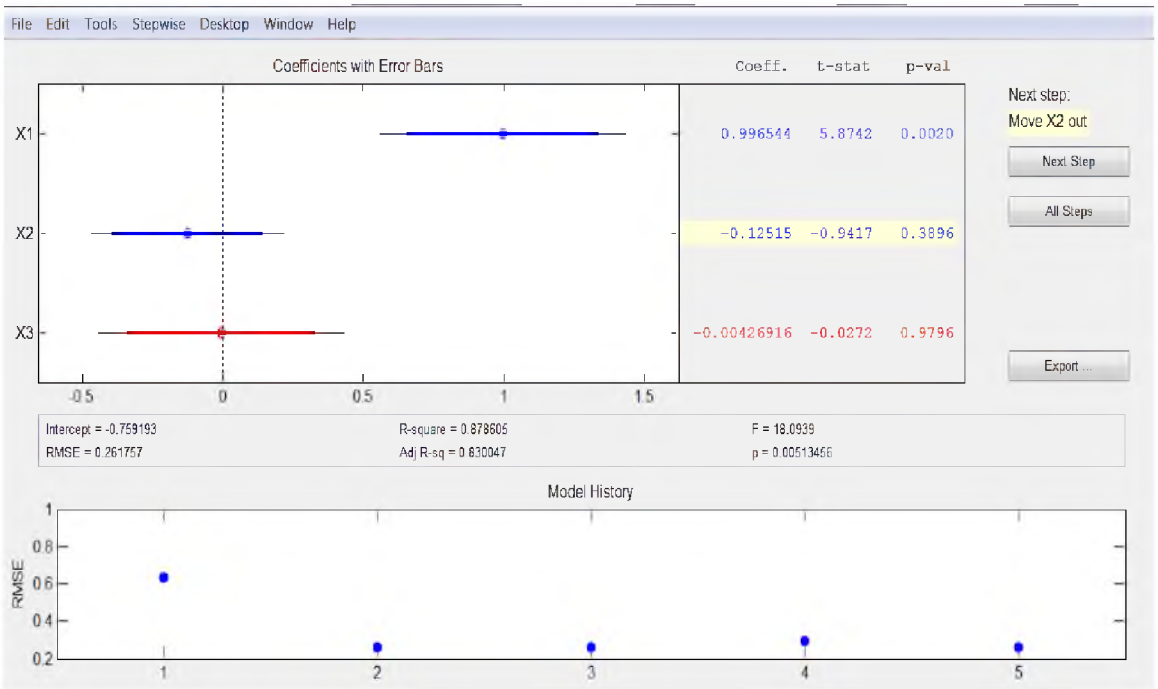


Figure 4.20. Second iteration in the benzaldehyde regression.

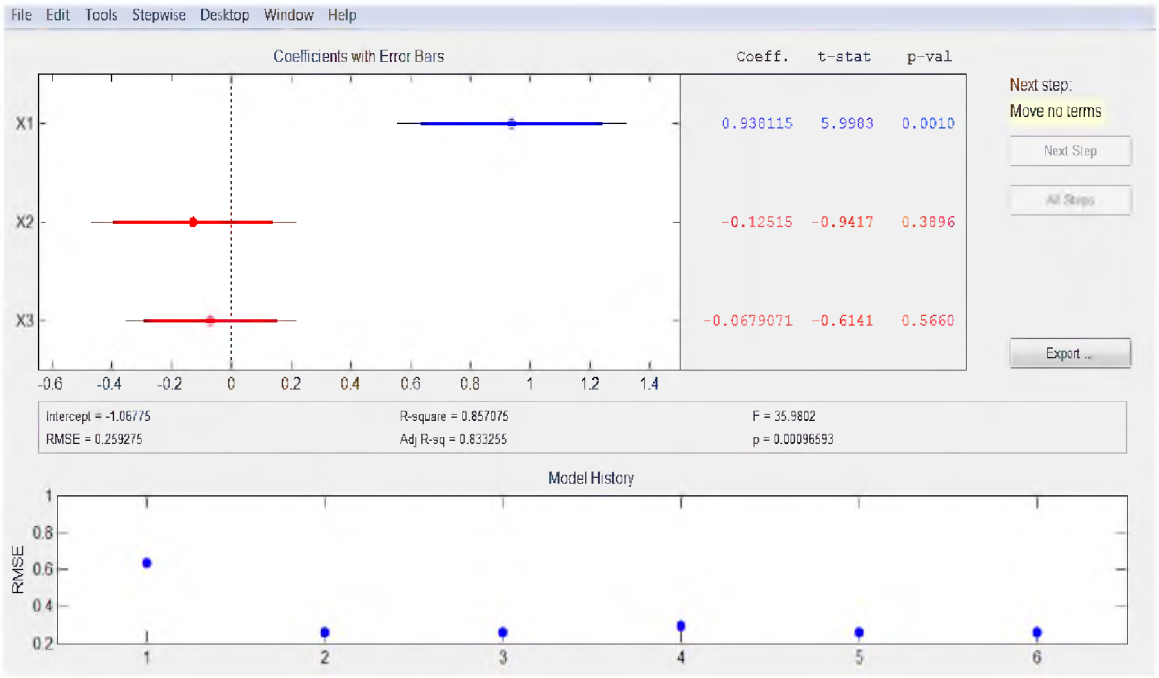


Figure 4.21. Final iteration for the benzaldehyde regression.

Table 4.4. Comparison of experimentally measured and predicted (from Eq. 1) enantioselectivities for the NHK allylation of benzaldehyde.

Measured $\Delta\Delta G^{\ddagger}$	***Predicted*** $\Delta\Delta G^{\ddagger}$
0.245	0.359
0.382	0.359
0.913	0.715
1.854	1.372
0.598	0.715
1.411	1.691
0.730	0.884
1.854	1.897

Table 4.5. Results of the LOO validation on the allylation of benzaldehyde data.

Substituent	***Measured*** $\Delta\Delta G^{\ddagger}$	***Predicted*** $\Delta\Delta G^{\ddagger}$
Me	0.245	0.40
Et	0.382	0.35
*i*Pr	0.913	0.98
*t*Bu	1.854	1.26
$CH(Pr)_2$	0.598	0.73
CEt_3	1.411	1.82
$CH(iPr)_2$	0.730	0.90
1-Ad	1.854	1.93

Model Development for the Allylation of Acetophenone

The acetophenone Sterimol model was derived using Matlab™ software. The student version, or a version which contains the statistics toolbox, is required to access the stepwise linear regression tool. The raw data were compiled into numeric matrices X and Y according to ligand substituent parameters and measured response, respectively, as shown in Table 4.6. Matrices X and Y were input into the Matlab™ software using X = [Matrix X] and Y = [Matrix Y] notations. Then the stepwise solver was initiated using the command "stepwise(X,Y)". In Figure 4.22, note that all the terms are shown in blue, which means that all terms are included in the present model. The model history has four points for the separate inclusion of each of the three terms. The coefficients are estimated for the B_1, B_5 and L terms in the upper right hand box. These values are also subjected to t- and p-tests to measure their significance in the model. This shows that the B_5 term is insignificant and should be eliminated from the model (as prompted by the "Next Step" indicator. Figure 4.23 gives the final model after the exclusion of the B_5 and L term. There is significant improvement in the model statistics after exclusion both terms. The only remaining terms are the offset term and the B_1 term, which produce Equation 4.4. Using Equation 4.4 the enantioselectivities for each of the synthetic ligands were predicted and compared with the measured values (Table 4.7).

Cross Validation for the NHK Allylation of Acetophenone Model

To validate the model given in Equation 4.4, a simple cross validation was performed. All data collected for the reaction were used in the training set to develop Equation 4.4. To cross validate this model as a measure of its true predictive power, a LOO validation was performed. These models were constrained to fit a model using only the terms of the base

Table 4.6. Matrix format for the allylation of acetophenone data.

	Matrix X			Matrix Y
Substituent	B_1	B_5	L	**Average** $\Delta\Delta G^{\ddagger}$
Me	1.52	2.04	2.87	-0.650
Et	1.52	3.17	4.11	-0.680
*i*Pr	1.9	3.17	4.11	-0.300
*t*Bu	2.6	3.17	4.11	0.470
$CH(Pr)_2$	1.9	4.54	6.17	-0.300
CEt_3	2.94	4.18	4.92	0.280
$CH(iPr)_2$	2.08	4.19	4.12	0.120
1-Ad	3.16	3.49	6.17	0.566

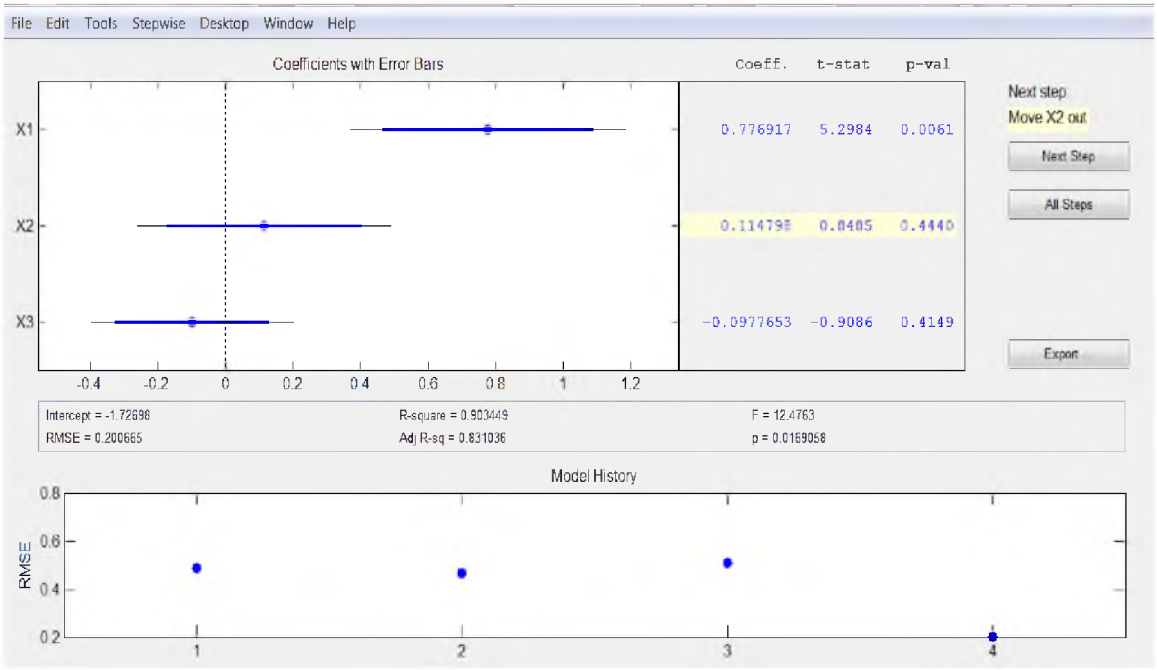


Figure 4.22. Initial acetophenone iteration.

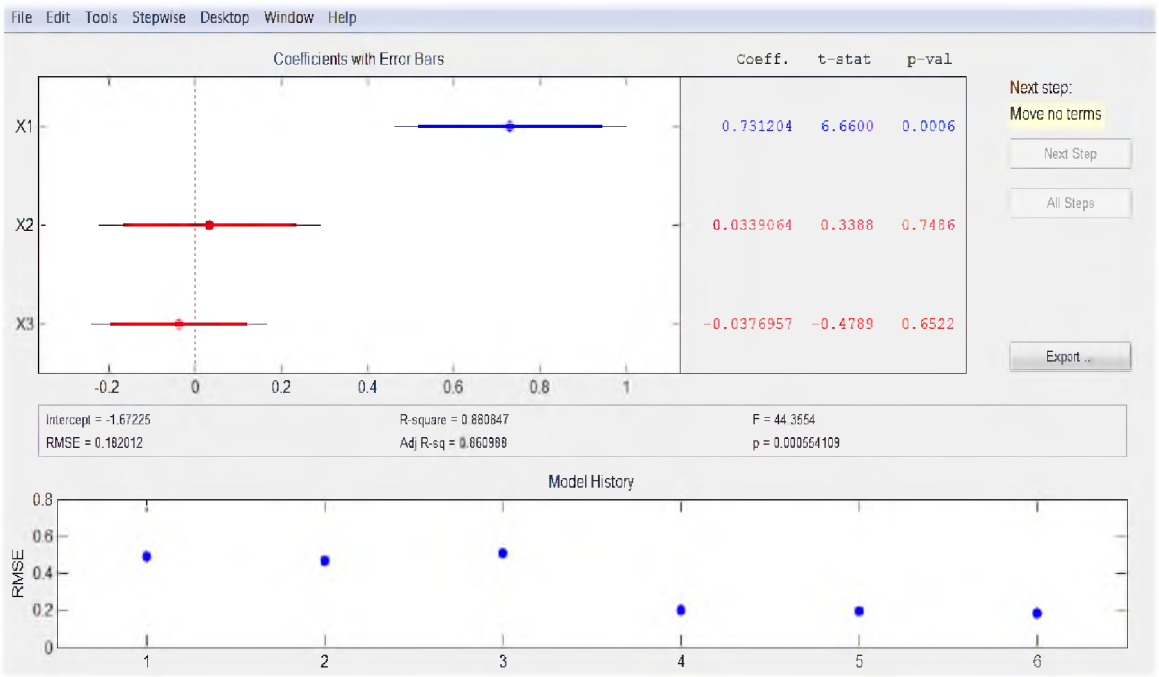


Figure 4.23. Final iteration in the modeling of the acetophenone data.

Table 4.7. Experimentally measured values compared to values calculated from Eq. 2 in the allylation of acetophenone.

Measured $\Delta\Delta G^{\ddagger}$	***Predicted*** $\Delta\Delta G^{\ddagger}$
-0.650	-0.560
-0.680	-0.560
-0.300	-0.283
0.470	0.228
-0.300	-0.283
0.280	0.476
0.120	-0.152
0.566	0.637

model of Equation 4.4 (z0, B_1) using regression techniques outlined above. These models were then used to independently predict the omitted data point as given in Table 4.8.

Model Determination for the Desymmetrization of Bisphenols

The desymmetrization Sterimol model was derived using Matlab™ software. The student version, or a version which contains the statistics toolbox, is required to access the stepwise linear regression tool. The raw data was compiled into numeric matrices X and Y according to ligand substituent parameters and measured response, respectively, as shown in Table 4.9. Matrices X and Y were input into the Matlab™ software using X = [Matrix X] and Y = [Matrix Y] notations. The stepwise solver was initiated using the command "stepwise(X,Y)". This command produced the prompt window shown in Figure 4.24. This prompt shows the coefficients for each term (X1, X2, and X3 correspond to B_1, B_5, and L, respectively) in the upper right box and also indicates whether the term is (blue) or is not (red) included in the current model. In Figure 4.24, note that all the terms are shown in blue, which means that all terms are included in the present model. The model history now has 4 points for the separate inclusion of each of the three terms. The coefficients are estimated for the B_1, B_5 and L terms in the upper right hand box. These values are subjected to t- and p-tests to measure their significance in the model at hand. These statistics show that the B_5 term is insignificant and should be eliminated from the model (as prompted by the "Next Step" indicator). Removal of this term leads to the model shown in Figure 4.25. Exclusion of the B_5 term leads to the model with the highest degree of statistical significance shown in Equation 4.5. Using this model, predictions were calculated for all the synthetic substrates, generating the data in Table 4.10.

Table 4.8. Results of the LOO-validation on the allylation of acetophenone data.

Substituent	Measured $\Delta\Delta G^{\ddagger}$	Predicted $\Delta\Delta G^{\ddagger}$
Me	-0.65	-0.52
Et	-0.68	-0.51
*i*Pr	-0.30	-0.28
*t*Bu	0.47	0.18
$CH(Pr)_2$	-0.30	-0.28
CEt_3	0.28	0.57
$CH(iPr)_2$	0.12	-0.19
1-Ad	0.57	0.69

Table 4.9. Matrix format for the desymmetrization data.

	Matrix X			Matrix Y
Substituent	B_1	B_5	L	Average $\Delta\Delta G^{\ddagger}$
Me	1.52	2.04	2.87	0.543
Et	1.52	3.17	4.11	0.596
Ph	1.71	3.11	6.28	0.517
Bn	1.52	6.02	4.62	0.421
*i*Pr	1.9	3.17	4.11	0.895
*t*Bu	2.6	3.17	4.11	1.724
Cy	1.91	3.49	6.17	0.746
CH_2tBu	1.52	4.18	4.89	0.582
$CHEt_2$	2.13	4.01	4.72	1.150
CH_2iPr	1.52	4.45	4.92	0.333
$CHPh_2$	2.01	6.02	5.15	0.780
1-Ad	3.16	3.49	6.17	1.724

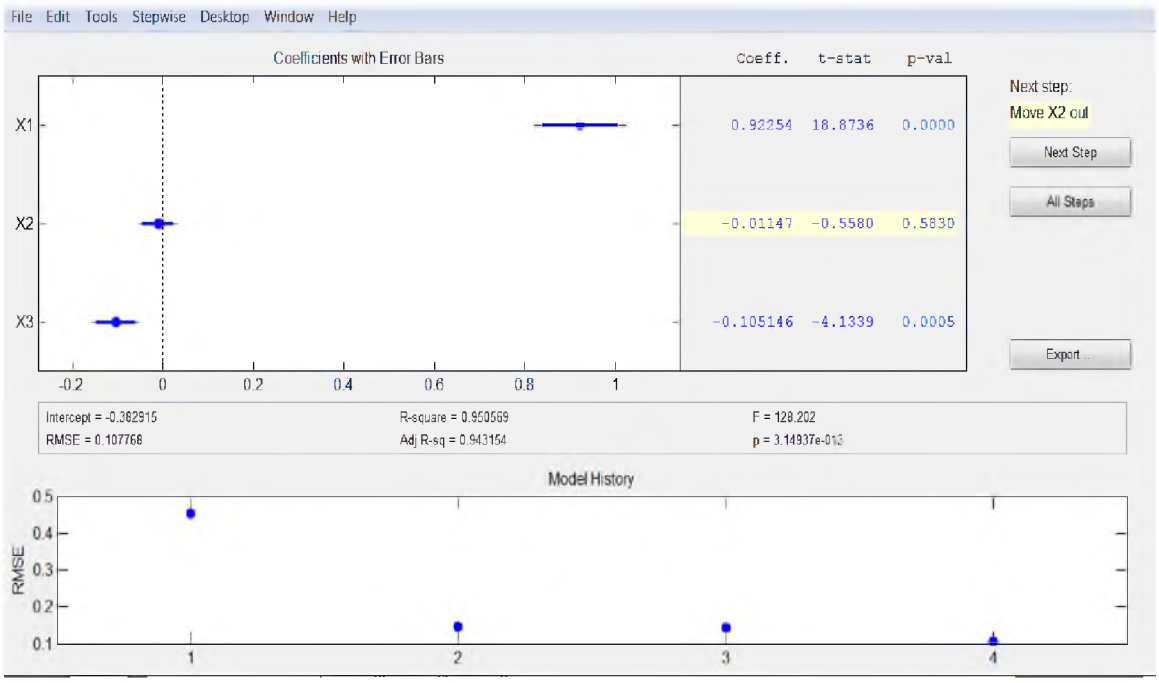


Figure 4.24. Initial inclusive model for the desymmetrization reaction.

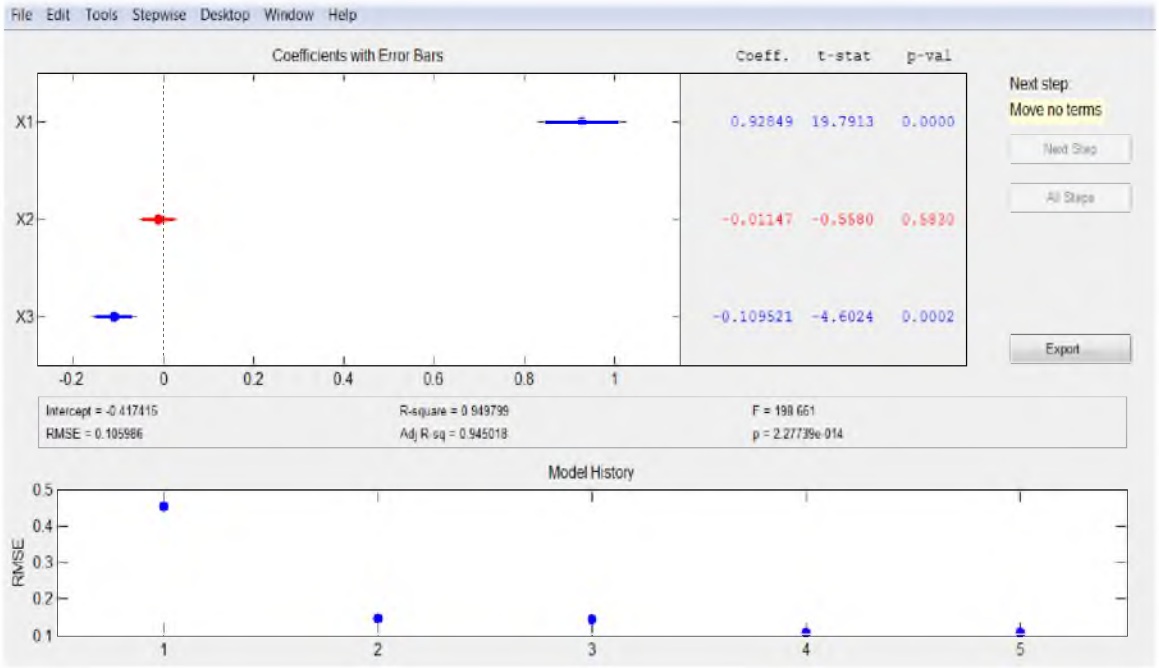


Figure 4.25. First and final iteration of the desymmetrization model.

Table 4.10. Experimentally measured values compared to those given by Eq. 3 for the desymmetrization reaction.

Measured $\Delta\Delta G^{\ddagger}$	***Predicted*** $\Delta\Delta G^{\ddagger}$
0.543	0.680
0.596	0.544
0.517	0.483
0.421	0.488
0.895	0.897
1.724	1.547
0.746	0.681
0.582	0.458
1.150	1.044
0.333	0.455
0.780	0.885
1.724	1.841

Cross Validation for the Desymmetrization of Bisphenol Model

To validate the model given in Equation 4.5, a simple cross validation was performed. All data collected for the reaction were used in the training set to develop Equation 4.5. To cross validate this model as a measure of its true predictive power, a LOO validation was performed. These models were constrained to fit a model using only the terms of the base model of Equation 4.5 (z0, B_1, L) using regression techniques outlined above. These models were then used to independently predict the omitted data point as given in Table 4.11.

Model Determination for the Propargylation of Acetophenone

The propargylation Sterimol model was derived using Matlab™ software. The student version, or a version which contains the statistics toolbox, is required to access the stepwise linear regression tool. The raw data was compiled into numeric matrices X and Y according to ligand substituent parameters and measured response, respectively, as shown in Table 4.12. Matrix X differs from previous like matrices because the data set contains variation in steric and electronic parameters. Thus, Matrix X includes crossterms between the electronic parameter and each of the Sterimol parameters. The crossterms are derived by multiplying the respective parameters for a given ligand. Matrices X and Y were input into the Matlab™ software using X = [Matrix X] and Y = [Matrix Y] notations. The stepwise solver was initiated using the command "stepwise(X,Y)". This command produced the prompt window shown in Figure 4.26. This prompt shows the coefficients for each term (X1, X2, X3, X4, X5, X6, and X7 correspond to B_1, B_5, L, σ, σB_1, σB_5, and σL, respectively) in the upper right box and indicates whether each term is (blue) or is not (red) included in the current model. Figure 4.26 shows that no terms are included into the model. The center box gives the intercept value as well as selected relevant statistics for the model given by the blue coefficients in the upper right box. With no terms

Table 4.11. Results of the LOO-validation on the desymmetrization of bisphenol data.

Substituent	Measured $\Delta\Delta G^{\ddagger}$	Predicted $\Delta\Delta G^{\ddagger}$
Me	0.543	0.783
Et	0.603	0.533
Ph	0.511	0.467
Bn	0.439	0.496
*i*Pr	0.885	0.898
*t*Bu	1.724	1.434
Cy	0.722	0.665
CH_2*t*Bu	0.582	0.437
$CHEt_2$	1.119	1.034
CH_2*i*Pr	0.349	0.474
$CHPh_2$	0.817	0.892
1-Ad	1.706	2.080

Table 4.12. Matrix formatted data for the propargylation of acetophenone.

		Matrix X							Matrix Y
Substituent		B_1	B_5	L	σ	σB_1	σB_5	σL	Measured $\Delta\Delta G^{\ddagger}$
E_{CF3}	S_{Me}	1.52	2.04	2.87	0.54	0.820	1.101	1.549	-0.127
E_{CF3}	S_{tBu}	2.6	3.17	4.11	0.54	1.404	1.711	2.219	0.0140
E_{CF3}	S_{CEt3}	2.94	4.18	4.92	0.54	1.587	2.257	2.656	0.237
E_{H}	S_{Me}	1.52	2.04	2.87	0	0	0	0	0.070
E_{H}	S_{tBu}	2.6	3.17	4.11	0	0	0	0	1.103
E_{H}	S_{CEt3}	2.94	4.18	4.92	0	0	0	0	0.688
E_{OMe}	S_{Me}	1.52	2.04	2.87	-0.27	-0.410	-0.550	-0.774	0.242
E_{OMe}	S_{tBu}	2.6	3.17	4.11	-0.27	-0.702	-0.855	-1.109	1.341
E_{OMe}	S_{CEt3}	2.94	4.18	4.92	-0.27	-0.793	-1.128	-1.328	0.808
E_{CF3}	S_{Me}	1.52	2.04	2.87	0.54	0.820	1.101	1.549	-0.070
E_{CF3}	S_{tBu}	2.6	3.17	4.11	0.54	1.404	1.711	2.219	0.286
E_{CF3}	S_{CEt3}	2.94	4.18	4.92	0.54	1.587	2.257	2.656	0.286
E_{H}	S_{Me}	1.52	2.04	2.87	0	0	0	0	0.117
E_{H}	S_{tBu}	2.6	3.17	4.11	0	0	0	0	1.140
E_{H}	S_{CEt3}	2.94	4.18	4.92	0	0	0	0	0.769
E_{OMe}	S_{Me}	1.52	2.04	2.87	-0.27	-0.410	-0.550	-0.774	0.237
E_{OMe}	S_{tBu}	2.6	3.17	4.11	-0.27	-0.702	-0.855	-1.109	1.430
E_{OMe}	S_{CEt3}	2.94	4.18	4.92	-0.27	-0.793	-1.128	-1.328	0.888

Figure 4.26. Initial regression command prompt.

included in the model, these statistics are meaningless in Figure 4.26. The bottom box is a history of the values of the root-mean-square error (RMSE) of each model calculated. By clicking on each term in the upper right hand box, each term is incorporated into the model to give the display shown in Figure 4.27. In Figure 4.27, note that all the terms are shown in blue, which means that all terms are included in the present model. The model history now has 8 points for the separate inclusion of each of the eight terms. The coefficients are estimated for the B_1, B_5, L and σ terms and their crossterms in the upper right hand box. These values are also subjected to t- and p-tests to measure their significance in the model at hand. This shows that the X7 term (σL) is insignificant and should be eliminated from the model (as prompted by the "Next Step" indicator). Removal of this term leads to the model shown in Figure 4.28. Figure 4.28 indicates that the X4 term (σ) should be eliminated, generating the model shown in Figure 4.29. Although this model appears adequate, through inspection it was determined that the model could be further simplified through elimination of the L term. Elimination of the L term gives an equally significant model shown in Figure 4.30. This model represents the most simplified statistically significant model and is shown in Equation 4.7. Equation 4.7 was used to predict values for all synthetic ligands, which were then compared to the measured values (Table 4.13).

Cross Validation for the NHK Propargylation of Acetophenone Model

To validate the model given in Equation 4.7, a similar cross validation was performed. All data collected for the reaction were used in the training set to develop Equation 4.7. To validate this model as a measure of its true predictive power, each individual data point was removed from the data set and the remaining data was used as a training set for a new model in a LOO validation. These models were constrained to fit a similar model (after showing inner-cross validation) using only the terms of the base model of Equation 4.7 (z0, B_1, B_5, E B_1, E B_5)

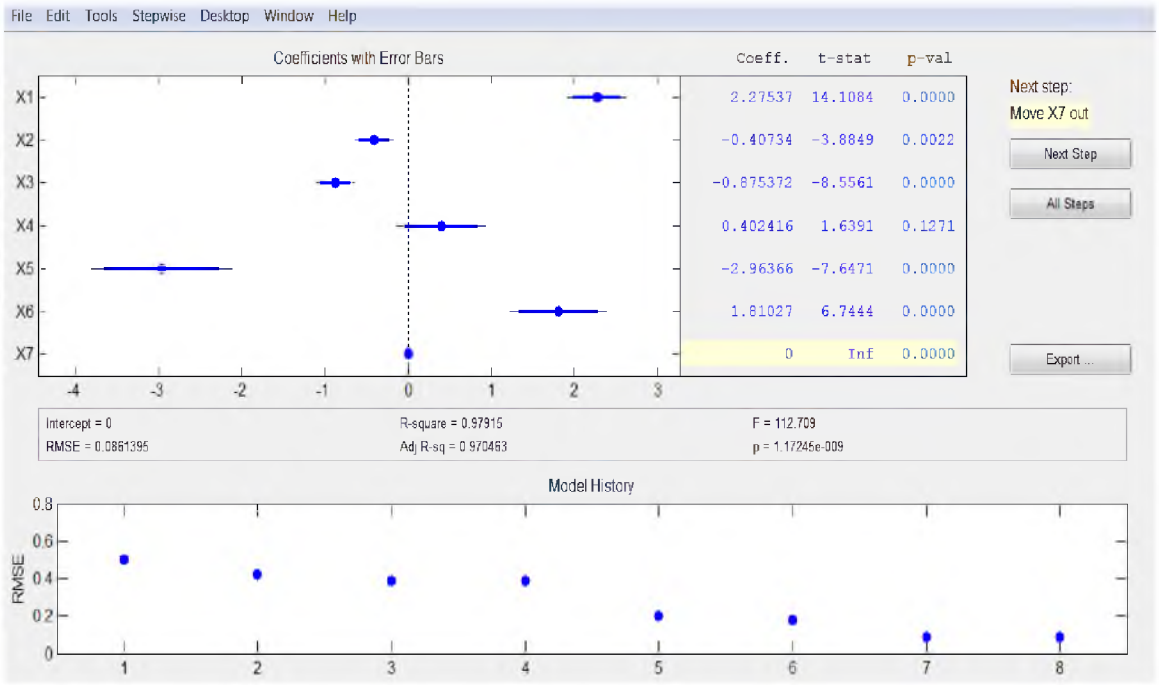


Figure 4.27. Initial incorporation of all variables in the regression.

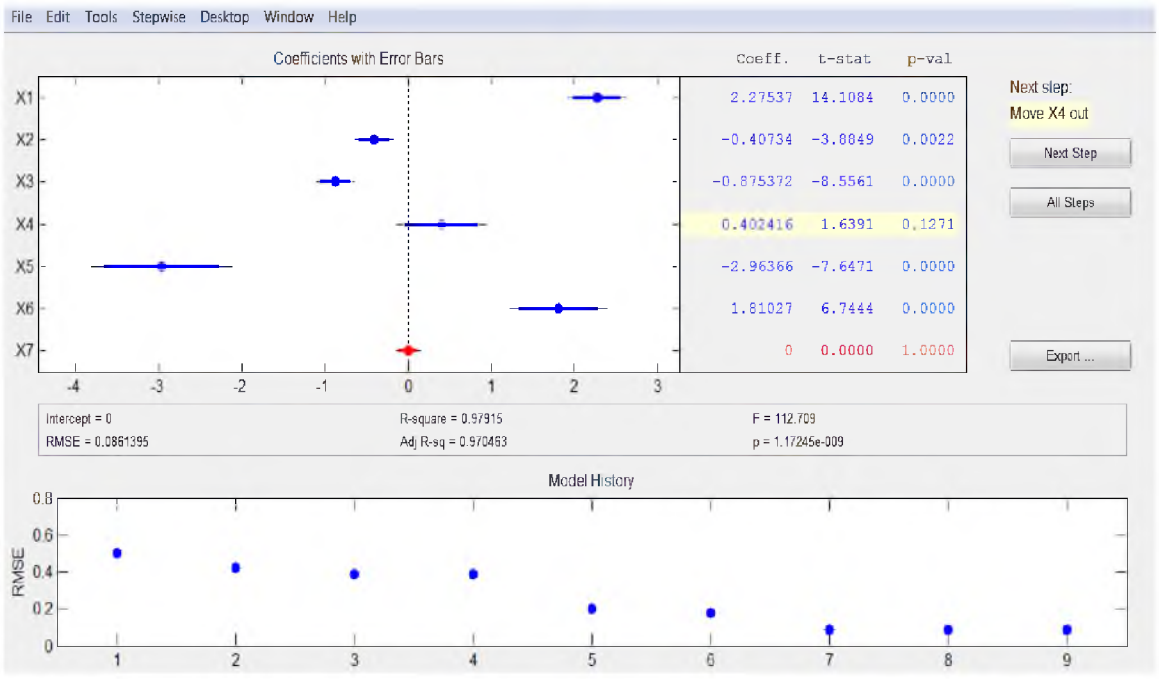


Figure 4.28. First iteration in the propargylation regression.

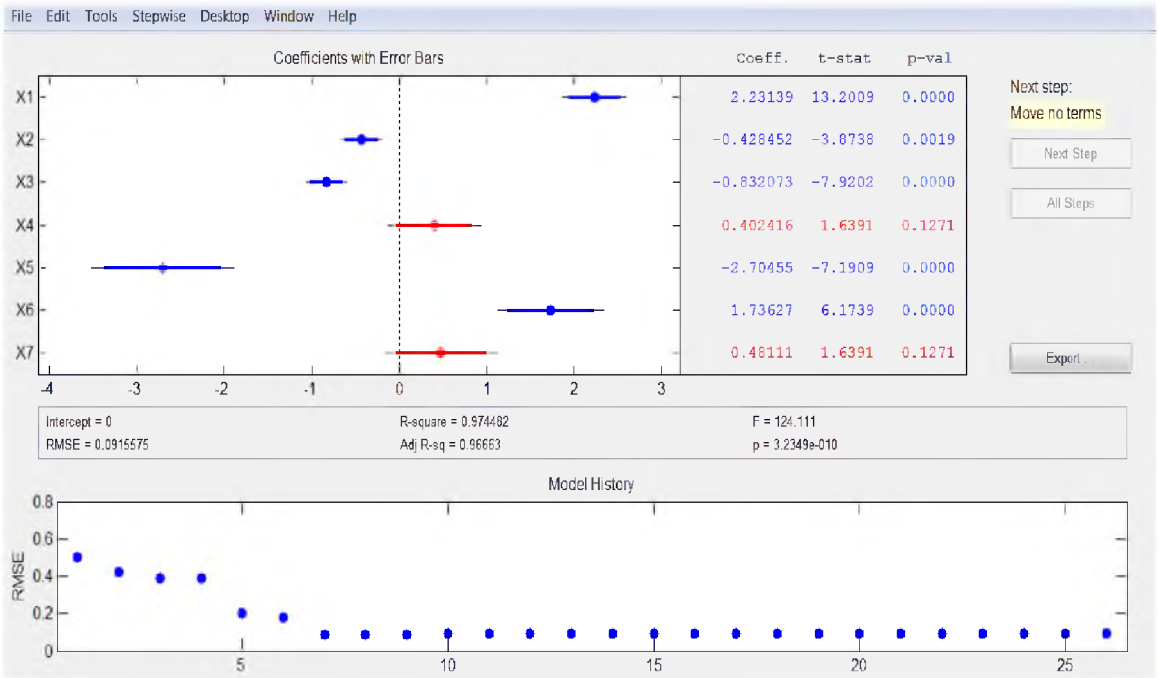


Figure 4.29. Second iteration in the propargylation regression.

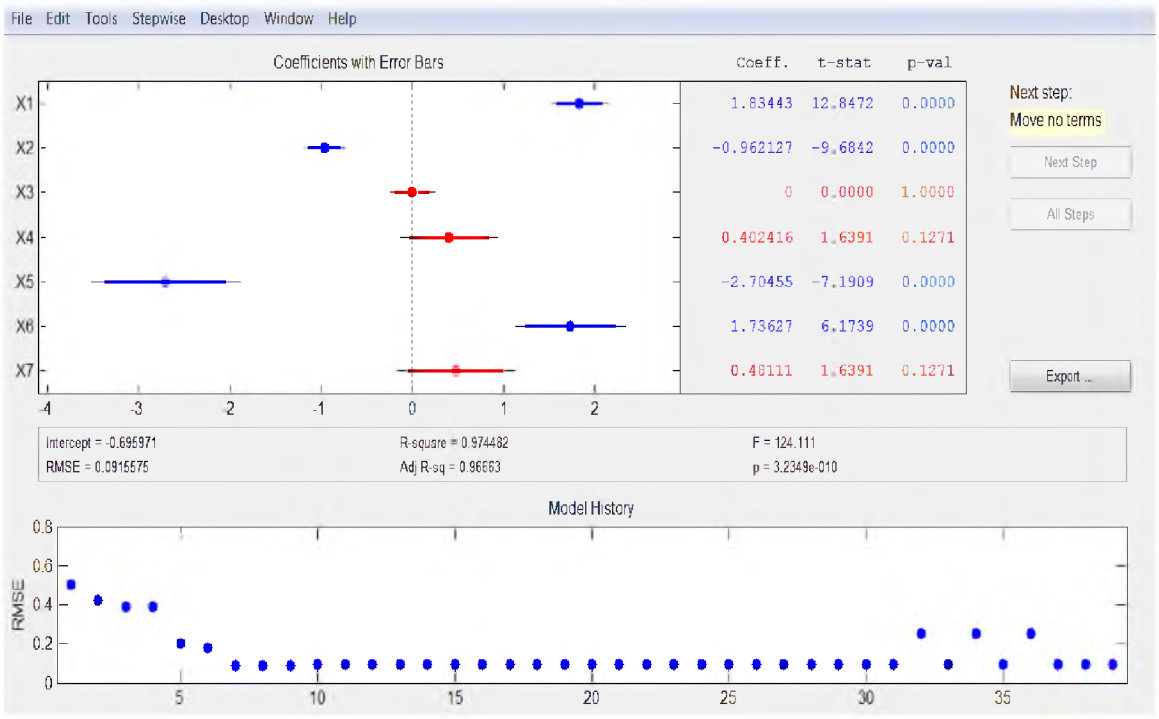


Figure 4.30. Final iteration of the propargylation model.

Table 4.13. Experimentally measured values compared to those given by Equation 7 for the desymmetrization reaction.

Measured $\Delta\Delta G^{\ddagger}$	***Predicted*** $\Delta\Delta G^{\ddagger}$
-0.127	-0.178
0.014	0.198
0.238	0.300
0.070	0.130
1.103	1.024
0.688	0.676
0.242	0.284
1.341	1.437
0.809	0.864

using regression techniques outlined above. These models were then used to independently predict the omitted data point as given in Table 4.14.

References

(1) Hansch, C.; Leo, A. *Exploring QSAR: Fundatmentals and Applications in Chemistry and Biology;* American Chemical Society: Washington, DC, 1995.

(2) Taft, R. W., Jr. *J. Am. Chem. Soc.* **1952**, *74*, 3120.

(3) Miller, J. J.; Sigman, M. S. *Angew. Chem., Int. Ed.* **2008**, *47*, 771.

(4) Sigman, M. S.; Miller, J. J. *J. Org. Chem.* **2009**, *74*, 7633.

(5) Harper, K. C.; Sigman, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 2179.

(6) Harper, K. C.; Sigman, M. S. *Science* **2011**, *333*, 1875.

(7) A. Verloop, J. T. In *Bological Activity and Chemical Structure*; Buisman, J. A., Ed.; Elsevier: Amsterdam, 1977, p 63.

(8) A. Verloop, J. T. In *QSAR in Drug Desing and Toxicology*; D. Hadzi, B. J.-B., Ed.; Elsevier: Amsterdam, 1987, p 97.

(9) Verloop, A. In *Drug Design*; Ariens, E. J., Ed.; Academic Press: New York, 1976; Vol. III, p 133.

(10) Fujita, T.; Takayama, C.; Nakajima, M. *J. Org. Chem.* **1973**, *38*, 1623.

(11) *Steric Effects in Organic Chemistry*; Newman, M. S., Ed.; Wiley: New York, 1956.

(12) Taft, R. W., Jr. *J. Am. Chem. Soc.* **1953**, *75*, 4538.

(13) Charton, M. *J. Org. Chem.* **1976**, *41*, 2217.

(14) Charton, M. *J. Am. Chem. Soc.* **1975**, *97*, 3691.

(15) Charton, M. *J. Am. Chem. Soc.* **1975**, *97*, 3694.

(16) Charton, M. *J. Am. Chem. Soc.* **1975**, *97*, 1552.

(17) Hansch, C.; Leo, A.; Taft, R. W. *Chem. Rev.* **1991**, *91*, 165.

(18) Winstein, S.; Holness, N. J. *J. Am. Chem. Soc.* **1955**, *77*, 5562.

Table 4.14. Results of the LOO-validation on the propargylation of acetophenone.

Substituent		Averaged Observed $\Delta\Delta G^{\ddagger}$	Predicted $\Delta\Delta G^{\ddagger}$
E_{CF3}	S_{Me}	-0.10	-0.23
E_{CF3}	S_{tBu}	0.15	0.64
E_{CF3}	S_{CEt3}	0.26	0.58
E_{H}	S_{Me}	0.09	0.15
E_{H}	S_{tBu}	1.12	0.97
E_{H}	S_{CEt3}	0.73	0.65
E_{OMe}	S_{Me}	0.24	0.31
E_{OMe}	S_{tBu}	1.39	1.55
E_{OMe}	S_{CEt3}	0.85	0.89

(19) Bott, G.; Field, L. D.; Sternhell, S. *J. Am. Chem. Soc.* **1980**, *102*, 5618.

(20) Adams, R.; Yuan, H. C. *Chem. Rev.* **1933**, *12*, 261.

(21) Tolman, C. A. *J. Am. Chem. Soc.* **1970**, *92*, 2956.

(22) Tolman, C. A. *Chem. Rev.* **1977**, *77*, 313.

(23) Verloop, A. In *IUPAC Pesticide Chemistry*; Miyamoto, J., Ed.; Pergamon: Oxford, 1983; Vol. 1, p 339.

(24) Werner, E. W.; Sigman, M. S. *J. Am. Chem. Soc.* **2010**, *132*, 13981.

(25) Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*; University Science Books: Sausalito, 2006.

(26) Swain, C. G.; Lupton, E. C. *J. Am. Chem. Soc.* **1968**, *90*, 4328.

(27) Swain, C. G.; Unger, S. H.; Rosenquist, N. R.; Swain, M. S. *J. Am. Chem. Soc.* **1983**, *105*, 492.

(28) Gustafson, J. L.; Sigman, M. S.; Miller, S. J. *Org. Lett.* **2010**, *12*, 2794.

(29) Lewis, C. A.; Gustafson, J. L.; Chiu, A.; Balsells, J.; Pollard, D.; Murry, J.; Reamer, R. A.; Hansen, K. B.; Miller, S. J. *J. Am. Chem. Soc.* **2008**, *130*, 16358.

(30) Palucki, M.; Finney, N. S.; Pospisil, P. J.; Güler, M. L.; Ishida, T.; Jacobsen, E. N. *J. Am. Chem. Soc.* **1998**, *120*, 948.

(31) Jensen, K. H.; Webb, J. D.; Sigman, M. S. *J. Am. Chem. Soc.* **2010**, *132*, 17471.